10/524094

P1 1063305

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

September 09, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/402,976*
FILING DATE: *August 13, 2002*
RELATED PCT APPLICATION NUMBER: *PCT/US03/25053*

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

T. LAWRENCE
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

08_14_02 Express Mail EU461948939US

Attorney Docket No. 00543-03

11051 U.S. PTO
08/13/02

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

Address to: Box Provisional Application
Assistant Commissioner for Patents
Washington, DC 20231
(703) 308-4357

Certificate Under 37 CFR 1.10
Date of Deposit: August 13, 2002
I hereby certify that this correspondence is being deposited with the United States Postal Service as "Express Mail" service under 37 CFR 1.10 on the date indicated above addressed to Box Provisional Application, Assistant Commissioner for Patents, Washington, DC 20231.

Sue Ann Carr

Express Mail No. EU461948939US

Jc973 U.S. PTO
60/402976
08/13/02

This is a request for filing a Provisional Application for Patent under 37 CFR 1.53(c)

Inventor(s) and Residence(s) (city and either state or foreign country):

Last Name	First Name	Middle Initial	City	State or Country
Kovatchev	Boris	P.	Amherst	Virginia
Cox	Daniel	J.	Charlottesville	Virginia

For: METHOD, SYSTEM, AND COMPUTER PROGRAM PRODUCT FOR PROCESSING OF SELF-MONITORING BLOOD GLUCOSE (SMBG) DATA TO ENHANCE DIABETIC SELF-MANAGEMENT

35 Sheets of specification.
_____ Sheets of drawings.

University of Virginia Patent Foundation claims small entity status as a nonprofit organization (37 CFR §1.9(e) and §1.27(d)). Therefore, please charge the Small Entity Fee of **$80** to **Deposit Account No. 50-0423.**

Please direct all communication relating to this application to:
Robert J. Decker, Esq.
Patent Counsel
University of Virginia Patent Foundation
1224 West Main Street, Suite 1-110
Charlottesville, VA 22903 U.S.A.

Telephone: (434) 924-2640
Fax: (434) 924-2493

This invention was made by an agency of the United States Government or under a contract with an agency of the United States Government. The government has certain rights in the invention.

YES ☐ NO ☒ Grant No. _____

Dated: August 13, 2002

Respectfully submitted,

By: _____
Robert J. Decker (Reg. No. 44,056)

METHOD, SYSTEM, AND COMPUTER PROGRAM PRODUCT FOR PROCESSING OF SELF-MONITORING BLOOD GLUCOSE (SMBG) DATA TO ENHANCE DIABETIC SELF-MANAGEMENT

CROSS-REFERENCES TO RELATED APPLICATIONS

The present invention is related to International Patent Application No. PCT/US01/09884 filed March 29, 2001 (Publication Nos. WO 01/72208 A2, WO 01/72208 A3), entitled "Method, System, and Computer Program Product for the Evaluation of Glycemic Control in Diabetes From Self-Monitoring Data," the entire disclosures of which are hereby incorporated by reference herein.

TABLE OF CONTENTS

OBJECTIVES

This phase 1 study provides, among other things, the continued development and the optimization of the mathematical models and algorithms first presented in the above-referenced PCT/US01/09884 Application. The study, but not limited thereto, validates and refines a new data analysis method, system and computer program product that comprises three algorithms for simultaneous evaluation, from routine SMBG data, of the two most important components of glycemic control in diabetes, HbA_{1c} and risk for hypoglycemia. This method pertains directly to enhancement of existing home BG monitoring devices by introducing an intelligent data interpretation component capable of predicting both HbA_{1c} and periods of increased risk for hypoglycemia. The data analysis method has three components (algorithms):

- **Algorithm 1:** Evaluation of HbA_{1c};
- **Algorithm 2:** Evaluation of long-term risk for severe hypoglycemia (SH), and
- **Algorithm 3:** Evaluation of short-term (within 24-48 hours) risk for hypoglycemia.

Algorithm 1 and 2 provide uninterrupted monitoring and information about the overall glycemic control of an individual with Type 1 or Type 2 diabetes mellitus (T1DM, T2DM), covering both the high and the low end of the BG scale. Algorithm 3 is supposed to be activated when Algorithm 2 indicates an increased long-term risk for hypoglycemia. Upon activation, Algorithm 3 requires more frequent monitoring (4 times a day) and provides 24 to 48-hour forecast of the risk for moderate/severe hypoglycemia.

Another important objective of Phase 1 was to test with existing data a number of hypotheses and ideas that could potentially lead to alternative algorithms estimating HbA_{1c} and computing risk for hypoglycemia in a manner that is conceptually different from the one proposed in the invention disclosure. The goal was to find potentially better solutions, or simply to verify that certain ideas do not lead to better results, which is essential for optimization and promptness of the analysis of the data that are currently being collected in Phase 2 of the study.

PROCEDURE

The funding for this Phase 1 study began on October 15, 2001 when an award for 40% of the contracted amount was issued to Boris Kovatchev. The award was used for a purchase of computers needed for the study and for partial salary support for Drs. Kovatchev and Cox, as planned. The study is supposed to continue 6 months and conclude with a report of its findings (this report). According to the contract between Lifescan and UVA, the rest of the funds will be released upon presentation of this report to Lifescan and will be used for continued salary support for Drs. Kovatchev and Cox, as well as to hire a programmer who will develop custom software implementing the refined algorithms. We have contacted the UVA Department of Systems Engineering and selected a graduate student who will develop the software code during the summer months of 2002. The software will be used for rapid analysis of the data that are currently collected by Phase 2 of this project and for eventual further implementation in a future Lifescan memory meter, or data processing software.

According to its objectives, the Phase 1 study continued the development and the optimization of the mathematical models and algorithms presented in our invention disclosure that is available to Lifescan. Throughout this text we assume knowledge of the definitions of various statistics computed from SMBG data, such as the Low and High BG Indices, BG rate of change, etc. that were presented in the invention disclosure. A detailed validation of these formulas with data for T1DM and T2DM patients is presented in a manuscript (*Kovatchev et al: Methods for Quantifying Self-Monitoring Blood Glucose Profiles Exemplified by an Examination of Blood Glucose Patterns in Patients with Type 1 and Type 2 Diabetes*) that was previously sent to Lifescan for review and is now submitted for publication to *Diabetes Technology and Therapeutics*.

DATA SETS

In order to ensure that the results of our optimization can be generalized to population level, Algorithms 1 and 2 were first optimized using training data sets and then tested for accuracy using an unrelated test data set. For Algorithm 3 we currently have only one data set containing parallel SMBG and records of SH. A detailed description of the patient population follows:

(1) Training Data set 1: Ninety-six patients with T1DM, who were diagnosed at least 2 years prior to the study. Forty-three of these patients reported at least two episodes of severe hypoglycemia in the past year and 53 patients reported no such episodes during the same

period. There were 38 males and 58 females. The mean age was 35±8 yr., mean duration of disease 16±10 yr., mean insulin units/kg per day 0.58±0.19, and mean HbA_{1c} was 8.6±1.8%. These subjects collected approximately 13,000 SMBG readings over 40-45-day period. The frequency of SMBG was approximately 3 reading/day. This data collection was followed by 6 months of monthly diaries of moderate and severe hypoglycemic episodes. This data set was used as a training data set for Algorithm 1 (no prior HbA_{1c}) and for Algorithm 2.

(2) Training Data set: Eighty-five patients with T1DM, diagnosed at least 2 years prior to the study, all of whom reported SH episodes in the past year. There were 44 males and 41 females. The mean age was 44±10 years, mean duration of disease 26±11 yr., mean insulin units/kg per day 0.6±0.2, the mean baseline HbA_{1c} was 7.7±1.1%, and the mean 6-month HbA_{1c} was 7.4±1% (6-month HbA_{1c} available for 60 subjects). These subjects collected approximately 75,500 SMBG readings during the 6 months between the two HbA_{1c} assays. The frequency of SMBG in Data set 2 was higher – 4-5 readings per day. In addition, during the 6 months of SMBG the subjects kept diaries of moderate and severe hypoglycemic episodes by date and time of their occurrence, resulting in 399 SH episodes. This data set was used as a training data set for Algorithm 1 (with prior HbA_{1c}) and for all analyses concerning Algorithm 3.

(3) The Test Data Set that we used contains data for N=600 subjects, 277 with T1DM and 323 with T2DM, all of whom used insulin to manage their diabetes. These data were collected by Amylin Pharmaceuticals, San Diego, CA and included 6-8 months of SMBG data (approximately 300,000 readings), accompanied by baseline and 6-month HbA_{1c} determinations and some demographic data. These subjects were participating in a clinical trial investigating the effects of pramlintide (in doses of 60 to 120 micrograms) on metabolic control. The subjects' use pramlintide was randomized across the T1DM and T2DM groups (Table 1).

Table 1: Demographic characteristics of the subjects in the Test Data Set.

Variable	T1DM - Mean(SD)	T2DM - Mean(SD)	p-level
Age (years)	38.0 (13.4)	58.1 (9.4)	< 0.001
Gender: Male/Female	136/141	157/166	Ns
Baseline HbA_{1c}	9.74 (1.3)	9.85 (1.3)	Ns
HbA_{1c} at month 6	8.77 (1.1)	8.98 (1.3)	0.04
Duration of diabetes (years)	14.6 (9.8)	13.5 (7.6)	Ns
Age at onset (years)	23.4 (12.8)	44.6 (10.4)	< 0.001
# SMBG readings / subject / day	3.2 (1.1)	2.9 (0.9)	< 0.005

Table 1 presents demographic characteristics and comparison of T1DM vs. T2DM subjects. For the first 6 months of the study the average HbA_{1c} declined significantly in both T1DM and T2 DM groups, perhaps due to the use of medication, which is out of the scope of this presentation (Table 1). This relatively rapid change in HbA_{1c} allowed for a better estimation of the predictive ability of Algorithm 1.

In all data sets SMBG was performed using Lifescan OneTouch II or OneTouch Profile meters.

ALGORITHM 1: EVALUATION OF HBA_{1c}

In the Phase 1 grant proposal our research goal was to optimize the prediction of HbA_{1c} (*Algorithm 1*) through: (1) Weighting higher the more proximal SMBG; (2) Weighting higher more prolonged high BG events: (3) Calibrating the High BG Index with an earlier HbA_{1c}, and (4) Incorporating other patient variables, such as age, gender and duration of disease.

Algorithm 1 includes an optimal function of SMBG data that evaluates subsequent HbA_{1c}, as well as recommendations for the optimal duration of the data collection period and the optimal frequency of self-monitoring during that period. It is essential to note, however, that the broader goal of Algorithm 1 is to evaluate the status of patients' glycemic control. Although HbA_{1c} is the accepted "gold standard" for evaluation of glycemic control, currently it is unclear whether another measure, such as average SMBG or High BG Index, would not be a better predictor of long-term complications in diabetes than HbA_{1c}. Until this issue is clarified, the goal of Algorithm 1 will be to estimate HbA_{1c}. In order to approximate as closely as possible future real applications of Algorithm 1 we proceeded as follows:

(1) First, several optimal functions using different independent variables, optimal duration, and optimal frequency of SMBG were derived from two *training data sets* 1 and 2, collected in our previous studies involving patients with T1DM;

(2) Then, all coefficients were fixed and Algorithm 1 was applied to the much larger *test data set* containing data for both T1DM and T2DM subjects collected under very different conditions in a clinical trial conducted by Amylin Pharmaceuticals.

(3) Detailed estimation of the preciseness of Algorithm 1 for various optimal functions was made using the *test data set only*.

This separation of *training* and *test* data sets allows us to claim that the estimated preciseness of Algorithm 1 can be generalized to any other data of subjects with T1DM or T2DM. Moreover, since the Amylin data (test data set) were collected from subjects who were undergoing treatment to lower their HbA_{1c}, and therefore exhibited unusually large variation of their HbA_{1c} over the 6-month period of observation, we can claim that Algorithm 1 is predictive not only of relatively constant HbA_{1c}, but also of large and unusually rapid changes in HbA_{1c}. Along these same lines, Algorithm 1 would be most useful for patients who have their goal to optimize their HbA_{1c}, which is presumably the patient group most likely to be interested in purchasing a meter with advanced features such as continuous evaluation of HbA_{1c}.

Summary of the Results

- The optimal SMBG data collection period is 45 days;
- The optimal frequency of SMBG is 3 reading per day;
- Two optimal HbA_{1c} estimating functions were developed: ***F1*** – using only SMBG data, and ***F2*** - using SMBG data plus an HbA_{1c} reading taken approximately 6 months prior to the HbA_{1c} that is being predicted;
- The evaluation of the accuracy of HbA_{1c} prediction in the *test data set* (N=573 subjects) was done by several criteria that are detailed in the following pages (Table 2). Here we will mention that in T1DM the overall accuracy (within 20% of measured HbA_{1c}) of ***F1*** was 96.5% and the overall accuracy of ***F2*** was 95.7%. For T2DM the overall accuracy of ***F1*** was 95.9%, the overall accuracy of ***F2*** was 98.4%. Thus, the accuracy of both ***F1*** and ***F2*** is comparable to a direct measurement of HbA_{1c};

- Most importantly, for patients whose HbA_{1c} changed 2 or more units from their baseline reading (N=68), the accuracy of ***F1*** in predicting this change was 100% in both T1DM and T2DM, while the accuracy of *F2* was 71% and 85% in T1DM and T2DM respectively;
- Both ***F1*** and ***F2*** provided substantially more accurate estimation of HbA_{1c} at 6 months than the original HbA_{1c} estimate at month 0. Using the average BG as a direct estimate of HbA_{1c} is not accurate as well;
- A number of alternative approaches were tested, such as selecting specific times of the day (postprandial reading) for evaluation of HbA_{1c}, different weighting of SMBG readings according to the elapsed time between each SMBG reading and HbA_{1c} determination, separate evaluation of subjects with different average blood glucose – to HbA_{1c} ratio, etc. While some of these alternative approaches achieved certain better results than the two functions proposed above, none was better overall. We can conclude that the optimal functions ***F1*** and ***F2*** will be used in future applications of Algorithm 1.

Detailed Results – *Test Data Set*

The most important part of the evaluation of Algorithm 1 is the evaluation of its performance on a data that are not related to the data used for its development and optimization. From the test data set, the data of 573 subjects, N=254 with T1DM and N=319 with T2DM, were complete enough to be used for the evaluation of Algorithm 1.

Optimal Algorithm 1: For each subject, a 45-day subset of his/her SMBG reading was selected. This subset had a starting date of approximately 75 days before the subject's 6-month HbA_{1c} assay and ending date approximately 30 days before that assay. Since in this data set the time of HbA_{1c} assay is known only approximately, the elapsed time between last SMBG reading taken into analysis and HbA_{1c} is not exact. This time period was selected through sequential optimization of its duration and its ending point (how long before HbA_{1c}). The optimal duration was 45 days. The optimal ending time was 1 month prior to HbA_{1c}. In other words, a 45-day SMBG would predict the values of HbA_{1c} approximately one month ahead. However, the prediction of any other HbA_{1c} value between days 45 and 75 is almost as good – the differences are numerical rather than of clinical significance. Similarly, the difference between a 45-day monitoring period and a 60-day monitoring period is not great. However, monitoring periods shorter than 45 days cause a rapid decline in predictive power.

The optimal estimation functions are linear and are given by the formulas:

Estimate 1 - Without prior knowledge of HbA_{1c}:

```
F1 = 0.809098*BGMM1 + 0.064540*LBGI1 - 0.151673*RHI1 + 1.873325
```

Estimate 2 - Knowing a prior HbA_{1c} (approximately 6 months ago).

```
F2 =0.682742*HBA0 + 0.054377*RHI1 + 1.553277
```

In these formulas `BGMM1` is the average blood glucose computed from the 45 days of SMBG readings; `LBGI1` and `RHI1` are the Low and the High BG Indices computed from the same readings, and `HBA0` is the baseline HbA_{1c} reading that is used for Estimate 2 only. The values of the coefficients are optimized using the training data set and relevant statistics and plots are presented in section Detailed Results – Training data set.

The functions ***F1*** and ***F2*** produce point estimates of HbA_{1c}, i.e. each function produces an estimated value of HbA_{1c}. Interval estimates can be obtained by using the regression error estimates presented in section Detailed Results – Training data set. However, applied to the test data set, these interval estimates will not be true 90% or 95% confidence intervals for HbA_{1c} because they are originally derived from the training data set and are only applied to the test data (see also the statistical note in the next section).

Evaluation of the accuracy of Algorithm 1: Tables 2A and 2B present results from the evaluation of the optimal Algorithm 1 with data from the test data set for subjects with T1DM and T2DM, respectively. Several criteria were used:

(1) Absolute deviation (AERR) of Estimated from measured HbA_{1c};
(2) Absolute percent deviation (PERR) of Estimated from measured HbA_{1c};
(3) Percent Estimates within 20% of measured HbA_{1c} (HIT 20),
(4) Percent readings within 10 of measured HbA_{1c} (HIT 10), and
(5) Percent readings outside of a 25%-zone around measured HbA_{1c} (MISS 25).

Table 2a: Accuracy of Algorithm 1 in T1DM (N=254 subjects).

	F1	F2	Average BG	Prior HbA_{1c}	P-value
AERR	0.77	0.61	1.68	1.1	< 0.001
PERR (%)	8.3	7.1	19.4	12.8	< 0.001
HIT20 (%)	96.5	95.7	61.0	81.0	< 0.001
HIT10 (%)	65.4	75.5	29.9	48.2	< 0.001
MISS 25 (%)	2.4	1.6	28.4	9.9	

Table 2b: Accuracy of Algorithm 1 in T2DM (N=319 subjects).

	F1	F2	Average BG	Prior HbA_{1c}	P-value
AERR	0.72	0.57	1.92	0.87	< 0.001
PERR (%)	7.6	6.4	20.9	11.7	< 0.001
HIT20 (%)	95.9	98.4	56.4	82.8	< 0.001
HIT10 (%)	70.2	79.3	29.5	53.3	< 0.001
MISS 25 (%)	1.2	0.6	36.7	8.2	< 0.001

The first two columns in Tables 2A and 2B present the results for the optimal functions ***F1*** and ***F2*** respectively. The third column presents the accuracy of the estimation if the average BG (in mmol/l) was taken as an estimate of HbA_{1c}. The fourth column presents the same accuracy measures computed using the HbA_{1c} assay at time 0 as an estimate of HbA_{1c} at 6 months. It is evident that for both T1DM and T2DM ***F2*** is a little better overall estimate of HbA_{1c} than ***F1***. Most importantly, both ***F1*** and ***F2*** are substantially better estimates of HbA_{1c} than its earlier value, or than the average BG. This is especially true for the % estimates that fell outside of the

25% accuracy zone. The difference between the performance of ***F1*** and ***F2*** and the estimate from a prior HbA_{1c} assay is highly significant (column 4).

Statistical Note: It is important to note that it is not appropriate to evaluate the accuracy of Algorithm 1 using traditional, regression-type criteria, such as R^2 or F and p values from ANOVA table. This is because the parameter estimates were derived from another unrelated data set (the training data) and are only applied to this test data set. Thus, statistical assumptions for the underlying model are violated (for example in the test data set the sum of the residuals will not be zero) and therefore R^2, F, and p lose their statistical meaning.

Further evaluation of the accuracy of Algorithm 1 in the test data set was done by reviewing the T1DM and T2DM subjects who had a substantial change in their SMBG reading from the baseline to 6-month follow-up. Tables 3A and 3B present list of the T1DM and T2DM subjects who had an absolute change in their HbA_{1c} equal to or greater than 2 units. In each subject group 34 subjects had such a change in HbA_{1c}. Algorithm 1, function F1, predicted 100% of such changes in both T1DM and T2DM. The predictive power of F2 was diminished due to the inclusion of baseline HbA_{1c} in the equation (which partially pulls the estimates back to the baseline value of HbA_{1c}) and was 71% in T1DM and 85% in T2DM. The baseline HbA_{1c} was outside of the 20% zone from 6-month HbA_{1c} for all but 2 subjects:

Table 3A: T1DM subjects who experienced change in their HbA_{1c} >= 2 units.

ID	HBA0	HBA6	DHBA	F1	F2	HIT F1	HIT F2	HIT HBA0
6504	12.0	7.0	5.00	6.82	9.90	100.00	.00	.00
6613	10.5	6.8	3.70	8.02	9.37	100.00	.00	.00
4003	12.4	8.9	3.50	8.45	10.73	100.00	.00	.00
6204	11.0	7.5	3.50	7.29	9.45	100.00	.00	.00
3709	13.0	9.7	3.30	8.99	11.54	100.00	100.00	.00
4701	12.8	9.5	3.30	9.50	11.61	100.00	.00	.00
3614	11.9	8.7	3.20	8.24	10.30	100.00	100.00	.00
3602	11.5	8.3	3.20	7.93	9.94	100.00	100.00	.00
6008	11.3	8.3	3.00	9.30	10.53	100.00	.00	.00
3723	13.0	10.1	2.90	8.80	11.46	100.00	100.00	.00
7010	12.7	9.8	2.90	8.09	10.89	100.00	100.00	.00
6208	11.5	8.7	2.80	8.42	10.09	100.00	100.00	.00
6202	10.6	7.8	2.80	7.91	9.37	100.00	.00	.00
3924	9.9	7.2	2.70	7.71	8.72	100.00	.00	.00
8211	11.0	8.3	2.70	8.76	10.32	100.00	.00	.00
6012	9.3	6.7	2.60	7.82	8.35	100.00	.00	.00
3913	11.0	8.4	2.60	7.88	9.54	100.00	100.00	.00
6701	11.2	8.6	2.60	8.75	10.07	100.00	100.00	.00
2307	10.6	8.1	2.50	7.95	9.27	100.00	100.00	.00
3516	11.8	9.3	2.50	7.76	10.03	100.00	100.00	.00
5808	9.6	7.2	2.40	7.61	8.52	100.00	100.00	.00
2201	11.8	9.5	2.30	8.90	10.71	100.00	100.00	.00
4010	12.4	10.1	2.30	8.57	11.15	100.00	100.00	.00
6210	11.9	9.6	2.30	8.33	10.40	100.00	100.00	.00
4904	11.3	9.1	2.20	8.63	10.29	100.00	100.00	.00
6709	10.3	8.1	2.20	7.83	9.04	100.00	100.00	.00
6619	9.5	7.3	2.20	7.64	8.57	100.00	100.00	.00
3921	10.9	8.8	2.10	7.20	9.19	100.00	100.00	.00
6603	11.0	8.9	2.10	8.18	9.89	100.00	100.00	.00
7415	10.6	8.5	2.10	7.94	9.27	100.00	100.00	.00
6515	9.8	7.8	2.00	7.13	8.54	100.00	100.00	.00
3611	10.3	8.3	2.00	8.36	9.23	100.00	100.00	.00

3732	13.2	11.2	2.00	9.30	11.99	100.00	100.00	100.00
7409	10.0	8.0	2.00	7.99	9.04	100.00	100.00	.00

Table 3B: T2DM subjects who experienced change in their HbA_{1c} >= 2 units.

ID	HBA0	HBA6	DHBA	F1	F2	HIT F1	HIT F2	HIT HBA0
6754	10.8	7.0	3.80	6.90	9.03	100.00	.00	.00
6361	11.3	7.6	3.70	8.51	10.20	100.00	.00	.00
6270	12.0	8.6	3.40	7.85	10.03	100.00	100.00	.00
6264	11.1	7.8	3.30	8.31	9.70	100.00	.00	.00
6355	11.8	8.6	3.20	7.99	9.90	100.00	100.00	.00
3961	10.8	8.0	2.80	9.13	9.73	100.00	.00	.00
6555	11.1	8.3	2.80	8.11	9.55	100.00	100.00	.00
8052	11.7	8.9	2.80	7.68	9.80	100.00	100.00	.00
5356	9.7	7.0	2.70	6.75	8.20	100.00	100.00	.00
3966	10.3	7.7	2.60	8.08	9.07	100.00	100.00	.00
908	9.5	6.9	2.60	7.47	8.23	100.00	100.00	.00
6554	10.7	8.1	2.60	8.16	9.42	100.00	100.00	.00
2353	11.1	8.7	2.40	8.99	9.90	100.00	100.00	.00
4064	11.3	8.9	2.40	7.89	9.88	100.00	100.00	.00
6351	10.1	7.7	2.40	7.92	8.63	100.00	100.00	.00
7551	12.2	9.8	2.40	9.17	11.02	100.00	100.00	.00
6358	8.4	6.1	2.30	7.00	7.32	100.00	.00	.00
3965	10.1	7.8	2.30	7.83	8.64	100.00	100.00	.00
914	11.1	8.8	2.30	9.57	10.33	100.00	100.00	.00
1603	10.2	7.9	2.30	8.02	8.88	100.00	100.00	.00
1708	10.8	8.6	2.20	7.62	9.24	100.00	100.00	.00
3761	12.4	10.2	2.20	9.13	10.86	100.00	100.00	.00
3768	11.2	9.0	2.20	8.29	9.74	100.00	100.00	.00
326	10.3	8.2	2.10	7.45	8.78	100.00	100.00	.00
109	9.3	7.2	2.10	7.70	8.18	100.00	100.00	.00
1501	11.9	9.8	2.10	8.52	10.18	100.00	100.00	.00
3964	13.7	11.6	2.10	10.08	12.65	100.00	100.00	100.00
4352	12.2	10.1	2.10	9.51	11.14	100.00	100.00	.00
7858	12.1	10.0	2.10	9.53	11.01	100.00	100.00	.00
4256	10.6	8.6	2.00	8.76	9.69	100.00	100.00	.00
4752	10.1	8.1	2.00	8.51	8.87	100.00	100.00	.00
6556	11.1	9.1	2.00	8.72	9.68	100.00	100.00	.00
6562	7.9	5.9	2.00	7.07	7.04	100.00	100.00	.00
8255	10.9	8.9	2.00	8.90	9.87	100.00	100.00	.00

In Tables 3A and 3B:

ID – subject's ID number;
HBA0 – baseline HbA_{1c};
HBA6 – measured 6-month HbA_{1c};
DHBA – absolute difference between baseline and 6-month HBA_{1c};
F1 – Estimated HbA_{1c} by Function ***F1***, SMBG data only;
F2 – Estimated HbA_{1c} by Function ***F2*** using prior HbA_{1c} assay;
Hit F1 = 100 if ***F1*** is within 20% of 6-month Hba_{1c} reading, 0 otherwise;
Hit F2 = 100 if ***F2*** is within 20% of 6-month Hba_{1c} reading, 0 otherwise, and
Hit HbA0 = 100 if baseline HbA_{1c} is within 20% of 6-month Hba_{1c} reading, 0 otherwise.

Detailed Results – *Training Data Set*

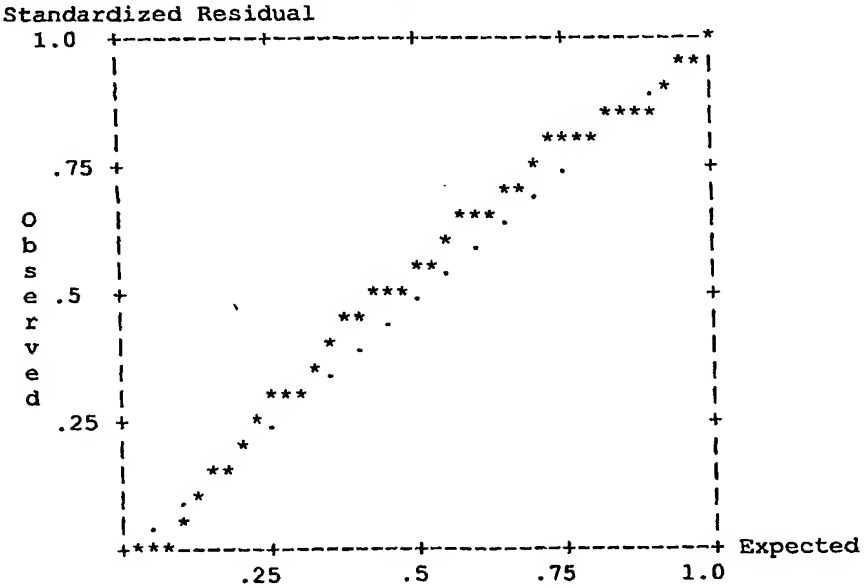
This section describes the steps to optimization of Algorithm 1. This optimization included two parts: (1) Assuming that no previous HbA_{1c} reading is available, and (2) Assuming that a prior HbA_{1c} could be used for prediction of HbA_{1c}.

Several different functions were considered for description of the relationship between SMBG data and HbA_{1c}. Optimal, in terms of accuracy and simplicity of computation, appeared to be a linear function of the average of SMBG readings, Low and High BG Indices, if no prior HbA1c reading is used and another linear function of a prior HbA_{1c} and the High BG Index. Nonlinear relationships did not enhance the goodness-of-fit of the models and therefore are not considered for practical application.

Training Data set 1 - No prior HbA_{1c}: A linear regression model was used to optimize the coefficient of function *F1*. The optimal coefficients were presented in the previous section. Here we give data about the goodness-of-fit of the model:

```
Multiple R            .71461
R Square              .51067

Analysis of Variance
                    DF        Sum of Squares      Mean Square
Regression           3             154.57097         51.52366
Residual            90             148.10903          1.64566

F =       31.30889       Signif F =  .0000
```

Analysis of the residuals of this model showed a close to normal distribution of the residuals (see the normal probability plot on the next page). The SD of the residuals was 1.2 (the mean is 0 by definition). Therefore we can accept that this model described the data well.



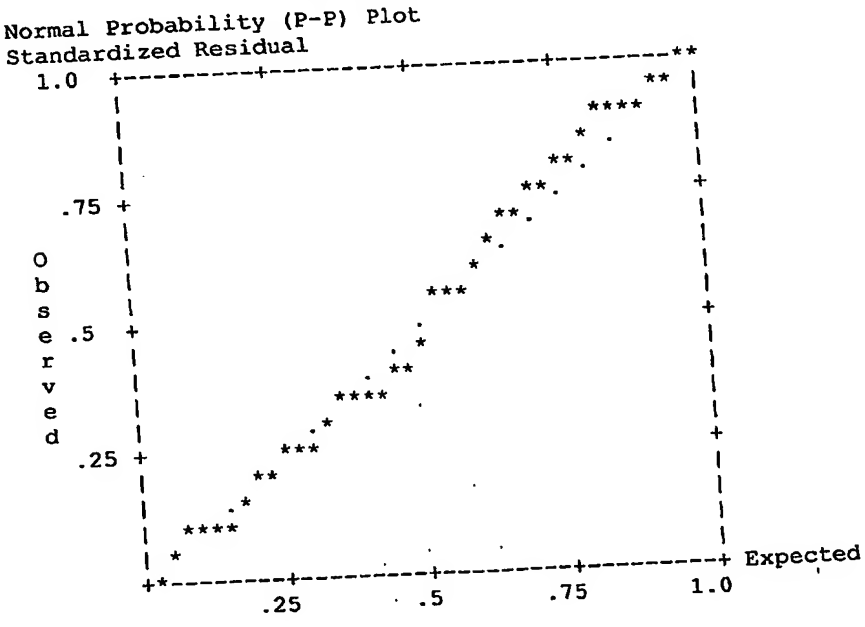
Training Data set 2 - Prior HbA_{1c}: Again, a linear regression model was used to optimize the coefficient of function *F2*. The optimal coefficients were presented in the previous section. Here we give data about the goodness-of-fit of the model:

```
Multiple R              .86907
R Square                .75528

Analysis of Variance
                DF        Sum of Squares        Mean Square
Regression       4              38.70237            9.67559
Residual        54              12.54000             .23222

F =       41.66522       Signif F =  .0000
```

Analysis of the residuals of this model showed a close to normal distribution of the residuals (see the normal probability plot on the next page). The SD of the residuals was 0.47. Therefore we can accept that this model described the data well.

In addition, comparing the models without and with a prior HbA1c, we can conclude that if a prior HbA1c is available for inclusion in the computations, the resulting model is substantially better both in terms of R^2 and in terms of residual error.

However, as we saw in the previous section, a prior HbA_{1c} does not contribute to the overall accuracy of prediction in an unrelated dataset and, in certain cases when HbA_{1c} changed substantially, is even obstructing the ability of the algorithm to account for rapid changes. Thus, we can conclude that, even if a prior HbA_{1c} maybe better from a statistical point of view, it may not have a sufficient practical utility to justify an inclusion of input of a reading in future meters. We also don't know what could be the elapsed time between an HbA1c assay and SMBG profile that would still render the HBA1c input useful. Perhaps this depends on the change of HBA1c during that time period – as we saw in the previous section, a change of 2 HbA1c units makes a prior HbA1c reading completely useless.



The Ratio of SMBG to HbA_{1c}

We will now present an alternative way to improve the statistical accuracy of the model fit and to keep a reasonable clinical applicability. It turns out that the ratio between the average of 45 days of SMBG readings and HbA_{1c} is a measure that has an almost perfect normal distribution (as evidenced by Kolmogorov-Smirnov test) and, most importantly, identifies three groups of subjects for whom this ratio is below 1.0, between 1.0 and 1.2, and above 1.2. Each of the first two groups accounts for approximately 40% of the subjects, the third group accounts for approximately 20% of the subjects. This is valid for both T1DM and T2DM and is observed in the training as well as in the test data sets. In addition, this ratio seems to be pretty stable over time and is perhaps a measure that reflects patients SMBG habits (for example, if SMBG is performed mostly at times when BG is low, the resulting average will underestimate HbA_{1c} and the corresponding ratio will be below 1.0). Keeping in mind that this is just a hypothesis that cannot be validated with the available data, we make some analyses that seem to demonstrate certain utility in knowing each person's ratio at some point of time. This may seem equivalent to knowing a prior HbA_{1c}, and it is perhaps equivalent in terms of data input, however the use of the ratio is very different than the use of a prior HbA_{1c}. Instead of being included directly into the prediction formula, the ratio is used to classify the person into using one of three different prediction formulas. These new formulas do not include HbA_{1c} directly and therefore do not suffer by the inertia of prediction that such inclusion may cause. In addition the average HbA_{1c} is not substantially different between the three groups defined by the ratio and is not correlated to the ratio, so the reason for different ratios in different persons must be unrelated to HbA_{1c}.

If we first classify the subjects in three groups according to their ratio and perform separate regression in the training data set, the goodness-of fit of the regression models increases substantially: (1) In group 1 (Ratio < 1.0) we get Multiple R=0.86 and R^2=0.73; (2) In group 2 (Ration between 1.0 and 1.2) the fit is almost perfect, R=0.97, R^2=0.94, and (3) In group 3 (ratio > 1.2) the fit is worst R=0.69, R^2=0.47. Since all three regression models do not include a prior HbA_{1c}, we can conclude that the goodness-of-fit increases dramatically for about 80% of the subjects, remains the same for the rest 20% of the subjects, and these subjects for whom the fit will be worse can be identified in advance.

Further, separating the test data set in three groups according to the subject's ratio, we get prediction accuracy similar to the accuracy we have achieved before (Table 4):

Table 4a: Accuracy of Algorithm 1 in T1DM (N=254 subjects).

	Ratio < 1.0	1.0 ≤ Ratio ≤ 1.2	Ratio > 1.2
AERR	0.70	0.63	0.74
PERR (%)	7.8	7.4	7.9
HIT20 (%)	93.8	93.0	95.5
HIT10 (%)	68.8	73.4	72.7
MISS 25 (%)	3.1	2.6	0.0

Table 4b: Accuracy of Algorithm 1 in T2DM (N=319 subjects).

	Ratio < 1.0	1.0 ≤ Ratio ≤ 1.2	Ratio > 1.2
AERR	0.63	0.68	0.89
PERR (%)	7.6	7.8	8.8
HIT20 (%)	97.4	95.0	95.3
HIT10 (%)	67.2	65.3	57.7
MISS 25 (%)	0.0	1.7	0.0

In short, knowing the SMBG-to HBA1c ratio for each subject and using separate estimates accordingly, seem to improve the statistical performance of the models without losing clinical accuracy.

Other hypotheses and ideas that were tested

We have tested a number of other hypotheses and ideas which may prove useful at least for prompt and more focused analysis of the data that are collected by Phase 2. A brief account of the results follows:

(1) HbA_{1c} is most associated (correlated) with SMBG readings taken in the afternoon hours – from 12 noon to 6 p.m. and least associated with fasting SMBG readings (4 a.m. - 8a.m.). However, it does not follow that taking only postprandial SMBG readings would improve the prediction of HbA_{1c}. on the contrary, the prediction would become worse if the [relatively small but important] contribution of all hours throughout the day is ignored. It is possible to improve a the prediction of HbA_{1c} a little if different hours throughout the day get different weighting, however the improvement is not sufficient to justify this additional complication of the model;

(2) The relationship between HbA_{1c} and average SMBG is substantially stronger in T2DM compared to T1DM, even if the two groups are matched by HBA_{1c}. In terms of direct correlation, in T1DM the coefficient is about 0.6 while in T2DM the coefficient is about 0.75 throughout the studies;

(3) Experiments with different weighting of SMBG reading dependent on the elapsed time between SMBG and HBA_{1c} assay (such as weighting higher more proximal results) did not yield better prediction of HBA_{1c};

(4) Inclusion of demographic variables, such as age, duration of diabetes, gender, etc., does not improve the prediction of HBA1c;

(5) The simplest possible linear relationship between HbA_{1c} and average SMBG (measured in mmol/l) is given by the formula: $HbA_{1c} = 0.41046*BGMM + 4.0775$. Although statistically inferior to the formulas ***F1*** and ***F2***, this formula provides HbA_{1c} estimates that are about 95% accurate in both T1DM and T2DM (in terms of deviation less than 20% from HbA_{1c} assay) and maybe useful if the computation of the Low and High BG Indices presents an issue for incorporation in a meter (however, the prediction of hypoglycemia cannot be done without computing the Low BG Index and therefore this formula might be useful only for meters that include Algorithm 1 but not include Algorithms 2 and 3).

ALGORITHM 2: EVALUATION of LONG-TERM RISK for SH.

In the Phase 1 grant proposal our research goal was to expand Algorithm 2 to include estimating individual probabilities for *biochemical significant hypoglycemia* (BSH, defined as BG reading <= 39 mg/dl) or *biochemical moderate hypoglycemia* (BMH, defines as 39 mg/dl < BG reading <= 55 mg/dl). In addition, we planned to evaluate whether Algorithm 2 predicts better occurrence of nocturnal (midnight to 7:00am) SH, compared to daytime SH.

Algorithm 2 is a classification algorithm. That is, based on SMBG data for a subject, it classifies the subject in a certain risk category for future BSH or MSH. In order to approximate as closely as possible future real applications of Algorithm 2 we proceeded as follows:

(4) First, several optimal classification variables and optimal classification categories optimal duration, and optimal frequency of SMBG were derived from *training data set 1*;

(5) Then, the test data set was split into two sections: first 45 days, and the rest of the data. The optimal parameters of Algorithm 2 were applied to the first 45-day portion of the data and the so estimated probabilities for future BSH or MSH were used to predict BSH and MSH in the second portion of the data;

(6) Detailed estimation of the preciseness of Algorithm 2, was made using *test data only*.

This separation of *training* and *test* data sets allows us to claim that the estimated preciseness of Algorithm 2 can be generalized to any other data of subjects with T1DM or T2DM. Moreover, since the Amylin data were collected from subjects who were undergoing intensive treatment, we can speculate that Algorithm 2 is tested and proven useful in subjects with changing and increasing risk for hypoglycemia.

Summary of the Results

- The optimal SMBG data collection period needed for estimation of the probability for future BSH or BMH is 40 to 45 days. The optimal frequency of SMBG is 3 to 4 readings per day. Larger number of readings does not lead to a substantial increase of the predictive power of Algorithm 2. With less than 3 reading per day the predictive power declines. However, this requirement refers to average number of readings per day for the 45-day observation period, it does not necessarily mean that 3-4 readings need to be performed every day;
- The relationship between predictor variables and future SH and MH is strictly nonlinear. Consequently, linear methods are not applicable for optimal prediction, although an R^2=50% can be achieved by a direct linear model (in comparison, the best result in the DCCT was 8% prediction of future SH);
- A separate prediction of nocturnal SH is generally weaker than prediction of daytime SH;
- Fifteen risk categories for future BSH and BMH were identified. The best separation of categories was achieved on the basis of the Low BG Index alone, although combinations between the low BG Index and other variables worked similarly well;
- Although the frequencies of BSH and BMH were different between T1DM and T2DM (see Table 5), the conditional frequencies, given a risk category, were not different between T1DM and T2DM. This allowed for unified approach to the risk of SH and MH;
- Various *empirical probabilities* for future were computed and compared for the 15 risk categories. All comparisons were highly significant, p's < 0.0005.
- These empirical probabilities were approximated by a two-parameter Weibull distributions yielding *theoretical probabilities* for future BSH and BMH in each risk category.
- The goodness-of-fit of these approximation was very good – all coefficients of determination were above 85%, some as high as 98% (see Figures 1-6).

Detailed Results – *Test Data Set*

Identifying personal risk categories for SH/MH: The data for all 600 subjects were used for these analyses. The Low BG Index (LBGI was computed for each subject from his/her first 45 days of SMBG data collection. Then, the LBGI for was classified in one of the 15 optimal risk categories (variable RCAT ranging from 0 to 14) as derived in training data set 1. These risk categories are defined by the inequalities:

```
if (LBGI le 0.25) RCAT=0.
if (LBGI gt 0.25 and LBGI le 0.5) RCAT=1.
if (LBGI gt 0.50 and LBGI le 0.75) RCAT=2.
if (LBGI gt 0.75 and LBGI le 1.00) RCAT=3.
if (LBGI gt 1.00 and LBGI le 1.25) RCAT=4.
if (LBGI gt 1.25 and LBGI le 1.50) RCAT=5.
if (LBGI gt 1.50 and LBGI le 1.75) RCAT=6.
if (LBGI gt 1.75 and LBGI le 2.00) RCAT=7.
if (LBGI gt 2.00 and LBGI le 2.50) RCAT=8.
if (LBGI gt 3.00 and LBGI le 3.50) RCAT=9.
if (LBGI gt 3.50 and LBGI le 4.00) RCAT=10.
if (LBGI gt 4.00 and LBGI le 4.50) RCAT=11.
if (LBGI gt 4.50 and LBGI le 5.25) RCAT=12.
if (LBGI gt 5.25 and LBGI le 6.50) RCAT=13.
if (LBGI gt 6.50) RCAT=14.
```

Observed frequency of BSH and BMH: For each subject, any occurrences of BSH and BMH registered by SMBG were counted for 1-monht, 3-month, and 6-month periods following the initial 45-day data collection. Table 5A presents the observed frequencies of 0, >=1, >=2, and >= 3 BSH and BMH for T1DM, Table 5B presents the same data for T2DM:

Table 5A: Observed frequency of BSH and BMH in T1DM

	BSH (BG ≤ 39 mg/dl)			BMH (39 mg/dl < BG ≤ 55 mg/dl)		
	1 month	3 months	6 months	1 month	3 months	6 months
Average # / Subject	0.82	1.77	2.74	3.64	8.33	12.93
% Ss with 0 episodes	62.8	50.8	46.6	25.2	18.0	17.7
% Ss with ≥ 2 episodes	18.8	33.1	38	64.3	75.6	77.1
% Ss with ≥ 3 episodes	9.8	23.3	28.2	50.8	68.0	71.1

Table 5B: Observed frequency of BSH and BMH in T2DM

	BSH (BG ≤ 39 mg/dl)			BMH (39 mg/dl < BG ≤ 55 mg/dl)		
	1 month	3 months	6 months	1 month	3 months	6 months
Average # / Subject	0.18	0.53	0.76	1.11	2.93	4.59
% Ss with 0 episodes	91.4	84.9	81.3	73.0	61.1	55.8
% Ss with ≥ 2 episodes	3.6	8.6	10.1	18.1	26.7	30.3
% Ss with ≥ 3 episodes	1.5	5.9	7.4	13.9	21.4	25.8

Nocturnal BSH and BMH represented approximately 15% of all episodes registered by SMBG. As in the training data set the correlation between nocturnal episodes and all predictor variables was weaker. We conclude that a targeted prediction of nocturnal episodes would be ineffective.

Empirical Probabilities for future BSH and BMH: Certain empirical probabilities for future BSH and BMH were computed in each of the 15 risk categories. These probabilities include: (1) Probabilities for at least one BSH or BMH within the next 1 month, 3 months, and 6 months; (2) Probabilities for at least two BSH or BMH within the next 3 months and 6 months, and (3) Probabilities for at least three BSH or BMH within the next 6 months. Of course, it is possible to compute any other combinations probabilities upon request.

A most important conclusion from this analysis was that, given a risk category, the probabilities for future BSH and BMH did not differ significantly between T1DM and T2DM. This allows for an unified approach to empirical and theoretical estimation of these probabilities in both T1DM and T2DM. Consequently, the data for T1DM and T2DM patients were combined for the following analyses.

Figures 1 through 6 (beginning on the next page) present scatter-plots of the six computed empirical probabilities plotted along the 15 risk categories. The empirical probabilities for BSH are presented by black triangles, while the empirical probabilities for BMH are presented by red squares.

All sets of empirical probabilities were compared across the 15 risk categories using univariate ANOVAs, and all p-levels were below 0.0005. Therefore, we observe highly significant differences between the frequencies of BSH and BMH episodes in the different risk categories.

Theoretical Probabilities for future BSH and BMH: In order to be able to use direct-formula estimation of the probabilities for future BSH and BMH, we approximated the empirical probabilities using two-parameter Weibull probability distribution. The Weibull distribution function is given by the formula:

$$F(x) = 1 - exp(-a.x^b) \text{ for any } x>0 \text{ and } 0 \text{ otherwise}$$

Statistical Note: The parameters a and b are greater than 0 and are called scale and shape parameter, respectively. In the special case b=1, Weibull's distribution becomes exponential. This distribution is frequently used in engineering problems as the distribution of randomly occurring technical failures that are not completely unrelated to each other (If the failures are completely unrelated, then they would form a Poisson process that would be described by an exponential distribution, e.g. b=1). The situation here is remotely similar – we need to describe the distribution of events (failures) that are not completely independent and tend to occur in clusters as evidenced by our previous research.

Each set of empirical probabilities was approximated by the theoretical formula given above. The parameters were estimated using nonlinear least squares (with initial parameter estimates given by a linear double-logarithmic model). The goodness-of-fit of each model was evaluated by its coefficient of determination (D^2). This statistics has a meaning similar to that of R^2 in linear regression, however R^2 is not applicable to non-linear models.

The model fits are presented in Figures 1 through 6 as black lines for the probabilities of BSH and as red lines for the probabilities of BMH. Above each figure we present the parameter estimates for the corresponding models, thus we give direct formulas for computing probabilities of 0, >=1, >=2, >=3 BSH or BMS episodes in periods of 1 month, 3 months, and 6 months following initial SMBG. Some of these formulas, or their versions, can be included in monitoring devices or software as indicators of risk for SH and MH.

The values of D^2 (and its square root D) are given below each figure as indicators of the preciseness of approximation. All values are above 85% and some reach 98%, which demonstrates that the approximation is very good and confirms that theoretical, instead of empirical probabilities could be used in future studies/application.

The theoretical probabilities for one or more moderate or severe hypoglycemic episodes are given by the formulas:

```
P(MH >= 1) = 1 - exp(-exp(-1.5839) * Risk** 1.0483)

P(SH >= 1) = 1 - exp(-exp(-4.1947) * Risk** 1.7472)
```

FIGURE 1: Probability of >= 1 Moderate or Severe Hypoglycemia within 1 Month

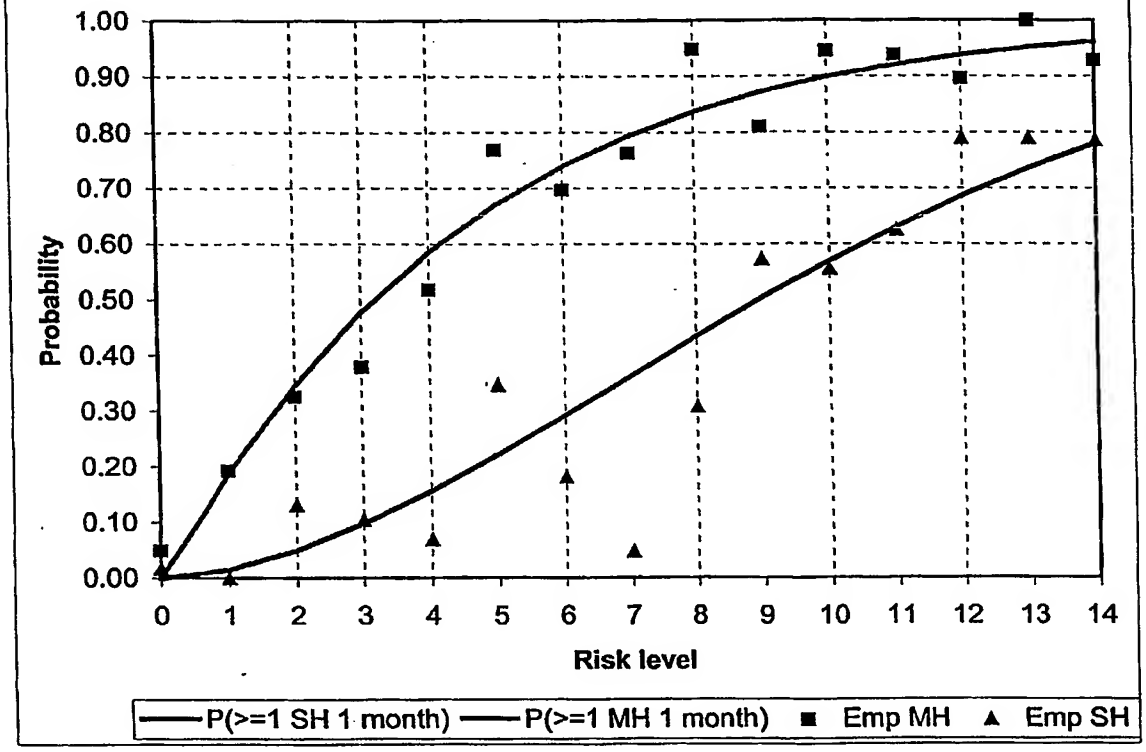


Figure 1 presents the empirical and theoretical probabilities for moderate (red line) and severe (black line) hypoglycemia within one month after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index. Since the models are nonlinear, the goodness-of-fit is evaluated by their coefficient of determination D^2, an analog of R^2 in linear models. The coefficients of determination and their square roots are as follows:

SH Model: **D^2 = 96%, D=98%.**

MH Model: **D^2 = 87%, D = 93%.**

The theoretical probabilities for one or more moderate or severe hypoglycemic episodes are given by the formulas:

```
P(MH >= 1) = 1 - exp(-exp(-1.3731) * Risk** 1.1351)

P(SH >= 1) = 1 - exp(-exp(-3.2802) * Risk** 1.5050)
```

FIGURE 2: Probability of >=1 Moderate or Severe Hypoglycemia within 3 Months

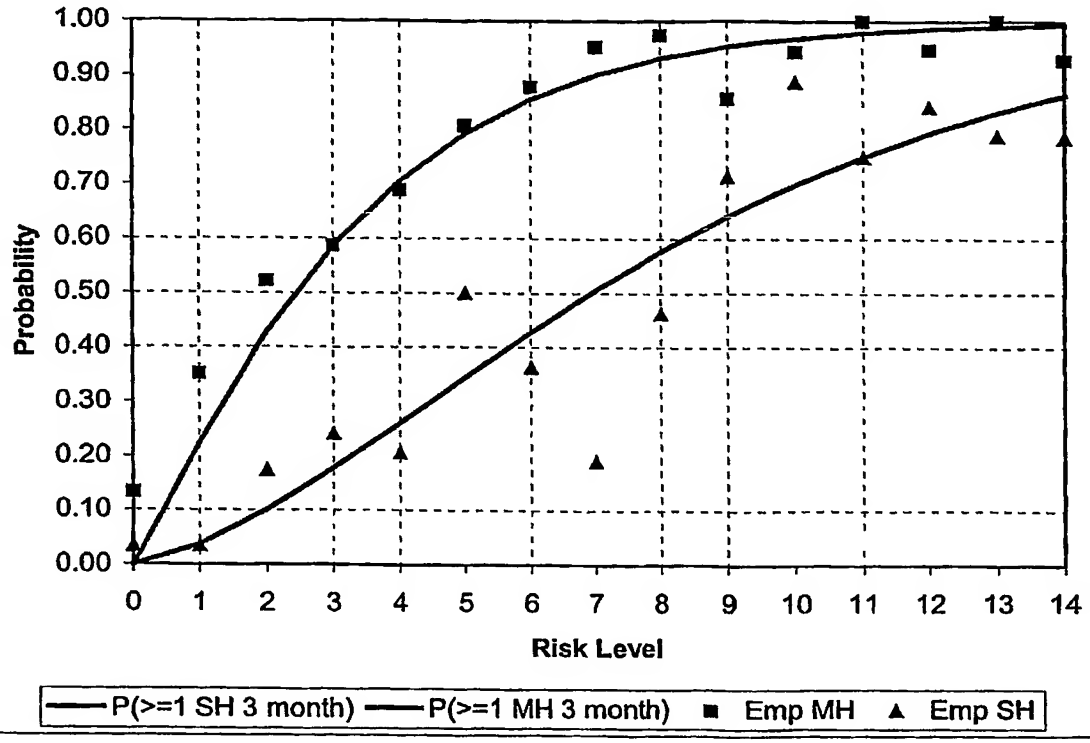


Figure 2 presents the empirical and theoretical probabilities for moderate (red line) and severe (black line) hypoglycemia within three months after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index.

The coefficients of determination and their square roots are as follows:

SH Model: **D^2 = 93%, D=97%.**

MH Model: **D^2 = 87%, D = 93%.**

The theoretical probabilities for one or more moderate or severe hypoglycemic episodes are given by the formulas:

```
P(MH >= 1) = 1 - exp(-exp(-1.3721) * Risk** 1.3511)

P(SH >= 1) = 1 - exp(-exp(-3.0591) * Risk** 1.4549)
```

FIGURE 3: Probability of >=1 Moderate or Severe Hypoglycemia within 6 Months

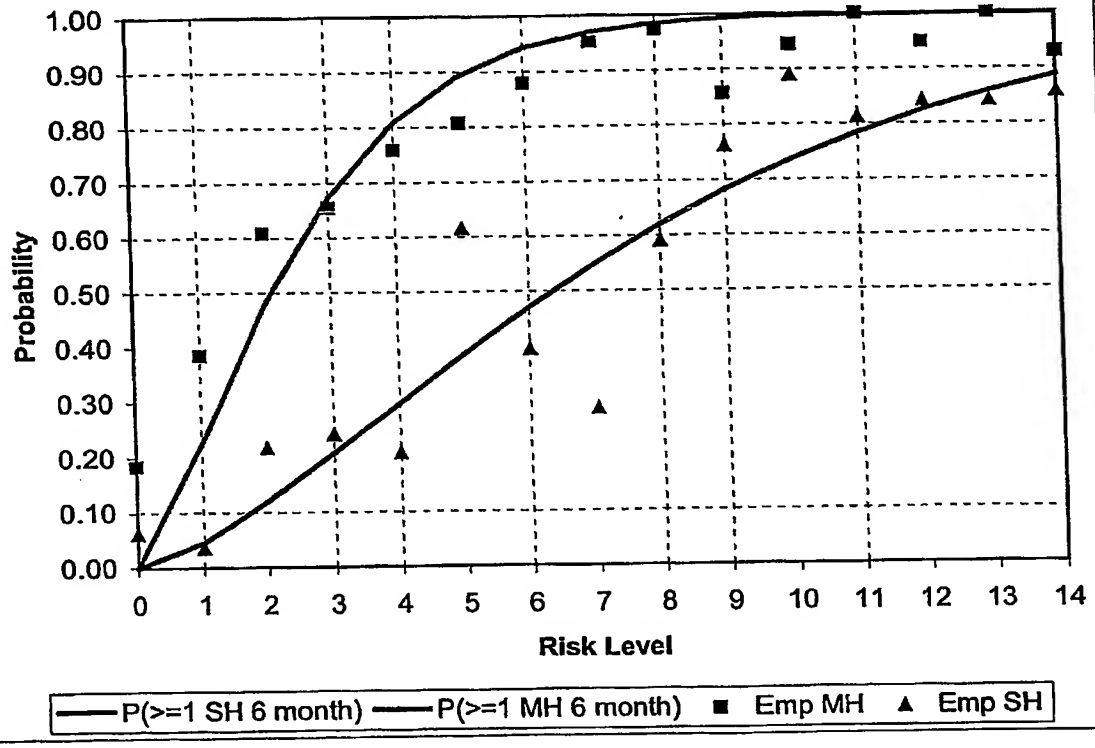


Figure 3 presents the empirical and theoretical probabilities for moderate (red line) and severe (black line) hypoglycemia within six months after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index.

The coefficients of determination and their square roots are as follows:

SH Model: D^2 = 86%, D=93%.

MH Model: D^2 = 89%, D = 95%.

The theoretical probabilities for two or more moderate or severe hypoglycemic episodes are given by the formulas:

```
P(MH >= 2) = 1 - exp(-exp(-1.6209) * Risk** 1.0515)

P(SH >= 2) = 1 - exp(-exp(-4.6862) * Risk** 1.8580)
```

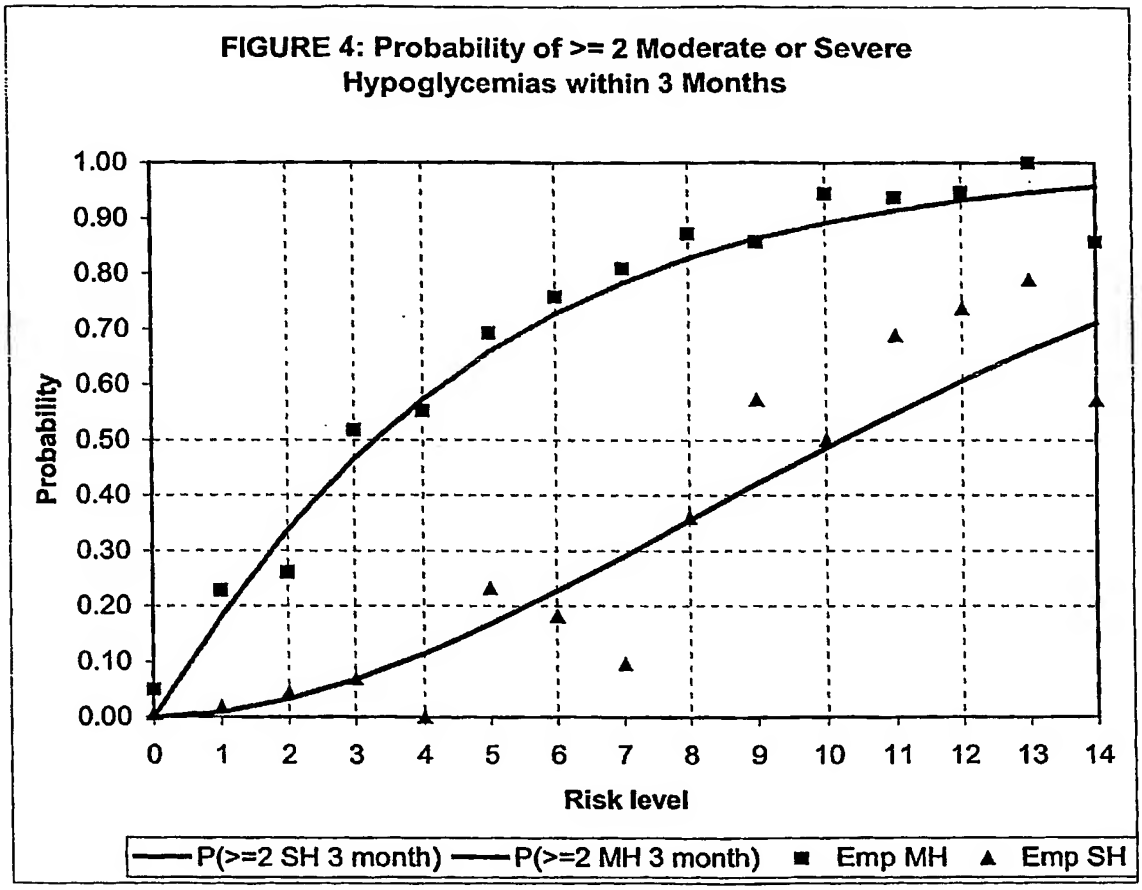


Figure 4 presents the empirical and theoretical probabilities for 2 or more moderate (red line) and severe (black line) hypoglycemic episodes within three months after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index.

The coefficients of determination and their square roots are as follows:

SH Model: **D^2 = 98%, D=99%.**

MH Model: **D^2 = 90%, D = 95%.**

The theoretical probabilities for two or more moderate or severe hypoglycemic episodes are given by the formulas:

$$P(MH >= 2) = 1 - \exp(-\exp(-1.7081) * Risk^{**} 1.1955)$$

$$P(SH >= 2) = 1 - \exp(-\exp(-4.5241) * Risk^{**} 1.9402)$$

FIGURE 5: Probability of >=2 Moderate or Severe Hypoglycemias within 6 Months

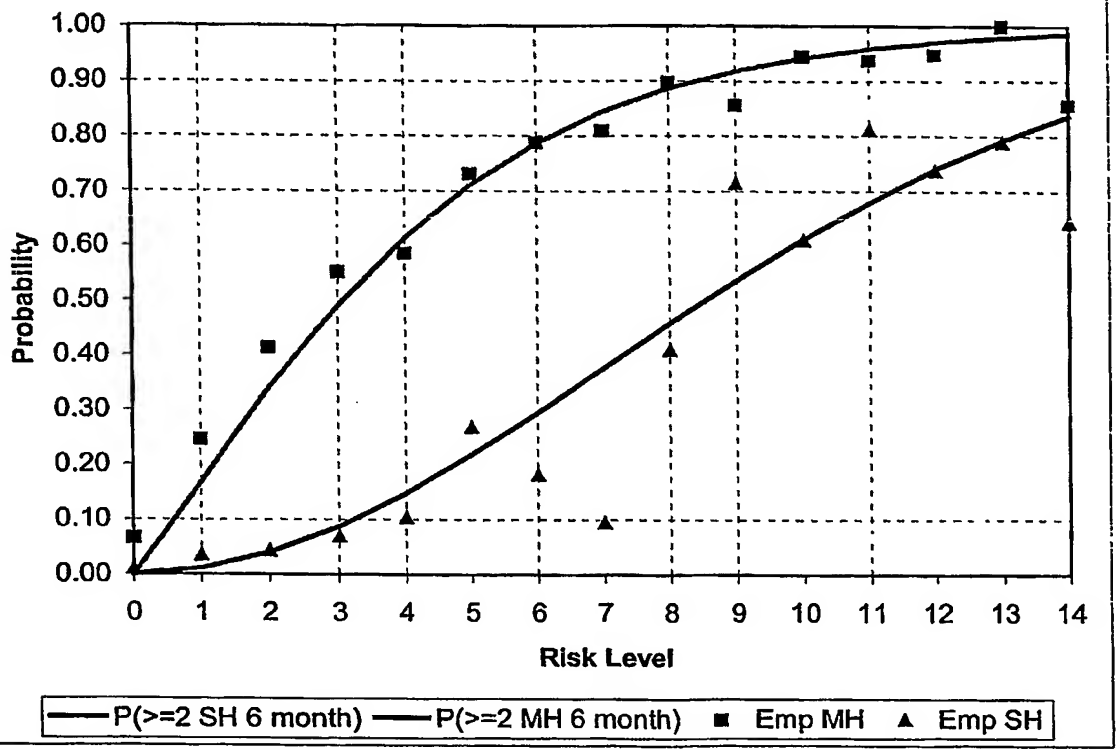


Figure 5 presents the empirical and theoretical probabilities for 2 or more moderate (red line) and severe (black line) hypoglycemic episodes within six months after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index.

The coefficients of determination and their square roots are as follows:

SH Model: **D^2 = 98%, D=99%.**

MH Model: **D^2 = 89%, D = 95%.**

The theoretical probabilities for three or more moderate or severe hypoglycemic episodes are given by the formulas:

$$P(MH >= 3) = 1 - \exp(-\exp(-2.0222) * Risk^{**} 1.2091)$$

$$P(SH >= 3) = 1 - \exp(-\exp(-5.5777) * Risk^{**} 2.2467)$$

FIGURE 6: Probability of >=3 Moderate or Severe Hypoglycemias within 6 Months

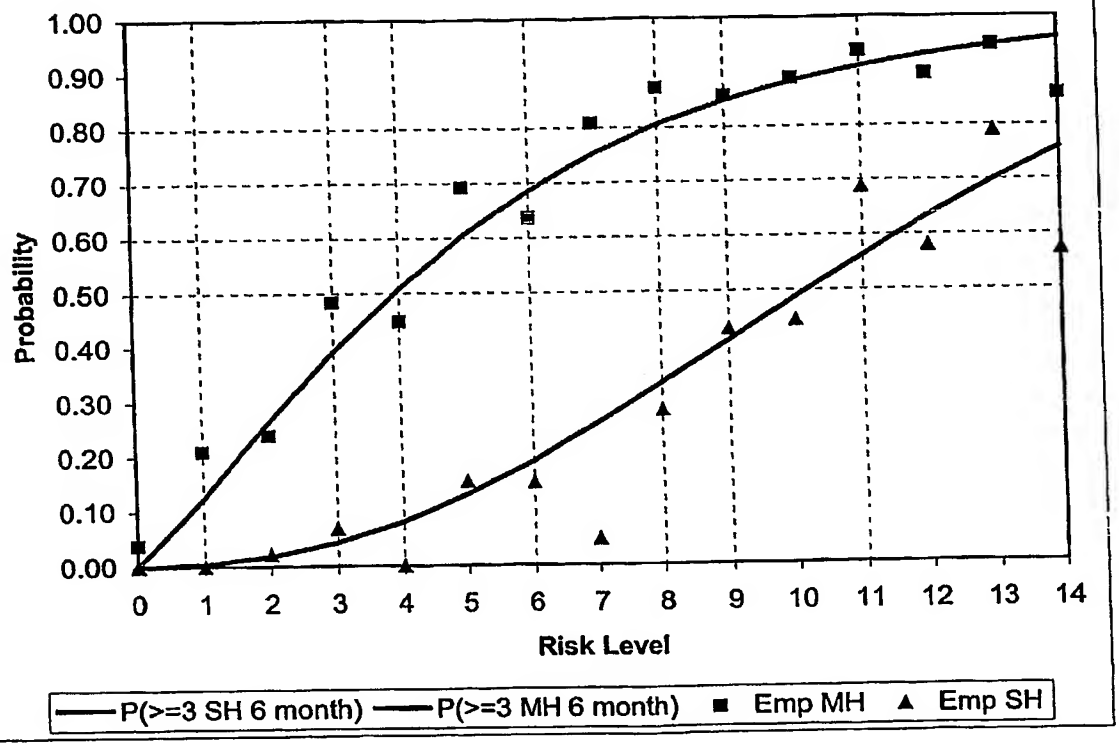


Figure 6 presents the empirical and theoretical probabilities for 3 or more moderate (red line) and severe (black line) hypoglycemic episodes within six months after the SMBG assessment for each of the 15 categories of risk level defined by the Low BG Index.

The coefficients of determination and their square roots are as follows:

SH Model: **D^2 = 97%, D=99%.**

MH Model: **D^2 = 90%, D = 95%.**

Detailed Results – *Training Data Set*

The training data set contained SMBG data followed by monthly diaries of severe hypoglycemia. As opposed to the test data set where BSH and BMH were identified by cutoff BG values, the monthly diaries contained report of symptomatic severe episodes defined as unconsciousness, stupor, inability for self-treatment, or significant cognitive impairment due to hypoglycemia. Within 6 months following SMBG the subjects reported on average 2.24 such episodes per person with 67% of the subjects reporting no such episodes. From a statistical point of view, this alone makes the distribution of SH episodes substantially skewed and unsuitable for application of linear methods. Nevertheless linear regression could be used to evaluate the relative contribution of various variables to the prediction of SH, but not for building the final model. We performed the following three analyses:

(1) No knowledge of SH history: Ignoring any knowledge of history of SH, we used regression to predict future SH from baseline HbA_{1c} and SMBG characteristics such as average BG, Low BG Index, and estimated BG risk rate of change (all variables are described in the original invention disclosure). As repeatedly found before, HbA_{1c} and average BG did not have any contribution to the forecast of SH. The final regression model included the Low BG Index and the BG risk rate of change and had the following goodness-of-fit:

```
Multiple R            .61548
R Square              .37882

Analysis of Variance

F =       27.74772       Signif F =  .0000

------------------ Variables in the Equation ------------------

Variable                B         SE B        Beta         T  Sig T

LBGI             4.173259      .649189    2.104085     6.428  .0000
RATE            -5.749637    1.091007   -1.724931    -5.270  .0000
(Constant)      -2.032859      .790491                -2.572  .0117
```

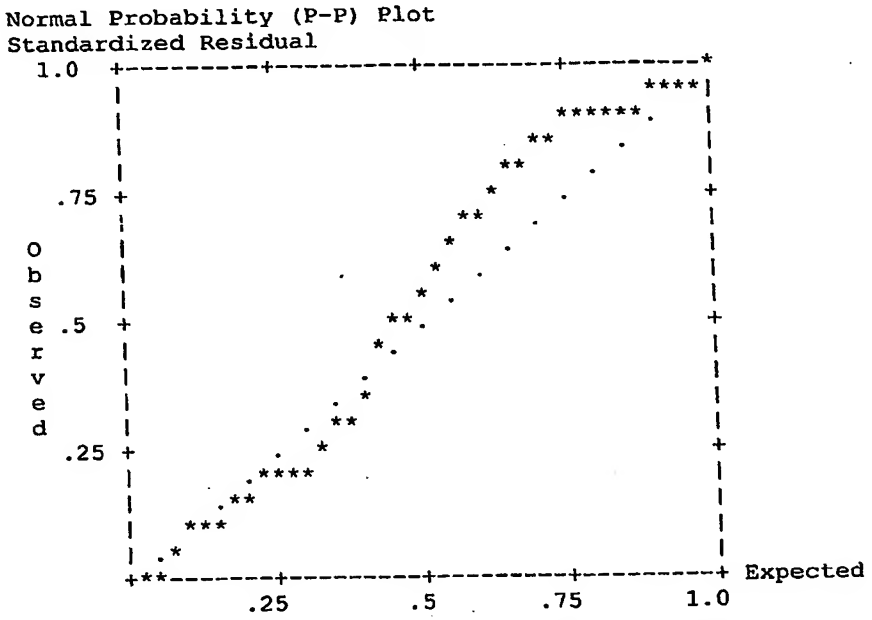
(2) Knowledge of prior SH: When we included the number of SH episodes in the previous year as reported in a screening questionnaire, this variable accounted for an additional 11% of the variance of future SH:

```
Multiple R            .70328
R Square              .49461

Analysis of Variance

F =       29.35999       Signif F =  .0000

------------------ Variables in the Equation ------------------

Variable                B         SE B        Beta         T  Sig T

SH               .337323      .074286     .375299     4.541  .0000
LDR            -4.350779    1.036380   -1.305264    -4.198  .0001
RLO             3.134519      .631684    1.580371     4.962  .0000
(Constant)     -2.136619      .717334                -2.979  .0037
```

(3) Without knowledge of the number of prior SH, just knowing whether a person had or did not have prior SH, we were able to account for 45% of the variance of future SH using only SMBG variables;

(4) Finally, two separate linear models accounted for 55% of the variance in daytime SH vs. 25% of the variance in nocturnal SH. The direct correlations of all predictor variables with nocturnal SH were also weaker. Nocturnal episodes represented 30% of all SH.

We conclude that a linear predictive model could directly account for about 40 to 50% of the variance of future SH. However, such a model is not well balanced in terms of its residual errors (which is due to the highly skewed distribution of the number of SH episodes across the diabetic population). A statistical evidence for that is given by the normal probability plot below, which shows a substantial deviation of the standardized residuals from their expected values:



Thus, we adopt another approach to predicting SH based on classification of subjects into risk categories using their SMBG data and estimation of the probabilities for subsequent SH in these categories. We attempted various classification models maximizing the difference between the risk categories and trying to achieve a maximum resolution of risk evaluation (in terms of maximal number of categories).

The best results were achieved by the classification based on the Low BG Index alone that had 15 risk categories (presented in the beginning of the previous section).

In addition to its best separation between categories, this result has other advantages as well: (1) No prior knowledge of history of SH is required; (2) The calculation is relatively simple and does not require tracking of temporal variables such as BG rate of change, and (3) The classification appeared to be equally applicable to both T1DM and T2DM patients (which is coherent with no requirements for knowledge of prior SH).

ALGORITHM 3: EVALUATION of SHORT-TERM RISK FOR HYPOGLYCEMIA

In the Phase 1 grant proposal our research goal was to optimize *Algorithm 3* in terms of:

(1) Utilization of baseline long-term risk (from *Algorithm 2*) and HbA_{1c} (from *Algorithm 1*);
(2) Risk criterion/threshold for hypoglycemia alert;
(3) Frequency of SMBG;
(4) Whether a hypoglycemia alert should be issued if an increased risk for hypoglycemia is detected and there is no SMBG for certain period of time, and
(5) Contribution of demographic variables such as history of severe hypoglycemia.

Introduction

As opposed to Algorithms 1 and 2, which have a longer history of development, Algorithm 3 deals with a proposition that was, until recently, considered impossible. In fact, there is still a general perception that prediction of any future BG value (hypoglycemia in particular) is not possible on the basis of previously known values (Bremer T and Gough DA. Is blood glucose predictable from previous values? A solicitation for data. *Diabetes*, 1999, 48: 445-451.). Our previous work, reported in one manuscript and presented in detail in the invention disclosure available to Lifescan, disputes this general perception. In order to explain the basis for this dispute and to clarify the reasoning behind Algorithm 3, we include the following paragraph.

Our "philosophy" in quantifying characteristics of diabetes: Hormonal interactions are governed by dynamic-control biochemical networks that have a more or less complex structure of principal nodes and conduits, depending on the studied endocrine system. Diabetes disrupts the network control of insulin-glucose dynamics at various levels. For example, in T1DM the natural production of insulin is completely eliminated, while in T2DM the utilization of insulin in the cell is obstructed by a greater insulin resistance. In T1DM (and frequently in T2DM) some form of external insulin replacement is required, which makes the control system vulnerable to imperfect external factors, including the timing and amount of pill or insulin injection, food eaten, physical activity, etc. This frequently leads to extreme BG excursions into hypoglycemia and hyperglycemia. In many, but not in all cases, hypoglycemia triggers an endocrine response, known as counterregulation. Thus, in mathematical terms, BG fluctuations over time are the measurable result of the action of a complex dynamic system, influenced by a number of internal and external factors. However, it is well known from the theory of dynamical systems that when the complexity of control increases, a purely deterministic system evolves to display random macro-behavior. Consequently, within short periods of time (minutes) the BG fluctuations observed at a human level would be *nearly-deterministic*, while over longer periods of time the fluctuations would be *nearly-random*, including extreme transitions, such as SH episodes. Thus, *stochastic modeling and statistical inference* are most appropriate for analysis of the system over longer periods of time – a paradigm adopted by Algorithms 1 and 2 that use our originally developed measures, such as the LBGI and HBGI, to predict, after a certain observation period, a range of values, or a probability of an event. Over short periods of time BG fluctuations can be modeled and predicted using *deterministic networks*, which would be the case with future intelligent insulin-delivery devices linked to continuous monitoring.

Algorithm 3 operates in an intermediate time scale of a few hours to a few days and therefore requires a *combination of statistical inference and deterministic modeling*. The former will be used to assess the baseline risk for SH for an individual, while the latter will be used for a dynamical tracking of individual parameters and forecast of SH episodes prior to their occurrence. When implemented in a device Algorithm 3 would work as follows:

(1) The device collects certain baseline information for the subject and establishes individual baseline parameters;

(2) Then, the device begins tracking a certain set of properties of the SMBG data;

(3) The device is equipped with a decision-making rule that decides when to raise a flag for upcoming SH and when to lower this flag if the data indicate that the threat is reduced;

(4) When the flag is raised, we assume that the subject is warned for SH in the following 24 hours (prediction time).

This dynamical prediction creates theoretical problems at both the level of model parameter optimization and at the level of evaluation of the preciseness of the optimal solution. We will begin with clarifying the second problem, as it is most important for understanding the action of Algorithm 3.

Evaluating the preciseness of Algorithm 3: While Algorithms 1 and 2 employ a static forecast and the criterion for evaluation of these algorithms is theoretically apparent – a better predictive value, with Algorithm 3 the optimization criterion is no longer straightforward. This is because by increasing the percentage of predicted SH episodes, we unavoidably increases the number of "raised flags," which in turn increases the number of potential "false alarms." The matter is additionally complicated by the fact that a "false alarm" is not clearly defined. In its pure form, a false alarm would be a raised flag that is not followed by and SH episode. However, SH could be avoided if the person perceives symptoms and takes an appropriate action. Thus, even if the biochemical potential for SH may present, an event may not occur. In order to deal with this problem we adopt the following optimization criterion:

(1) Maximize the prediction of upcoming SH within 24 hours;

(2) Minimize the ratio R_{ud} of duration periods of "flag up" to "flag down".

While the first of these two points is clear, the second may need an additional explanation. Looking from the perspective of an implementation of Algorithm 3 in a meter, at every SMBG determination the meter decides whether to raise a flag or not to raise a flag for upcoming SH. When the flag is raised, it may stay up for some time (along several subsequent SMBG readings) until a decision is made to take the flag down. Thus, we will have an alternating process of "flag up" and "flag down" with the changes happening at points of SMBG. The ratio R_{ud} referred to in point (2) above, is the average time for a person, counted while the flag is up, divided to the average time counted while the flag is down.

Our previous best result presented in the Invention disclosure was a prediction of 44% of SH episodes within 24 hours, and R_{ud}=1:7, e.g. one day of high-risk alert was alternating with 7 days of no alert. Since at that time we assumed that the warning period was at least 24 hours, the algorithm was optimized to raise a flag no more frequently than once a week. Given that this analysis was done using data for subjects who were experiencing high rate of SH episodes, this ratio was considered acceptable.

During Phase 1 of this study we had to use the same data set for refinement of Algorithm 3 since there is no other data available that include simultaneous SMBG records and records of SH. We also used a similar criterion to evaluate the preciseness of Algorithm 3. However, we *changed substantially everything else*. The tracking of the data, the parameter estimation, all threshold values and the decision-making rule are no longer the same. These changes were caused by a new idea that SH is preceded by certain "depletion" of the body's reserves to counterregulate and that this *depletion can be tracked by using SMBG data*. The exact implementation of this idea is described in the section "Decision-making Rule." Since the decision-making rule involves a continuous criterion and a somewhat artificial cutoff, several solutions are presented and one is selected as optimal for further investigation. However, upon presentation of these results, we may decide to select another solution to be implemented in future applications of Algorithm 3.

Summary of the Results

First, it is important to note that all results presented below go well beyond statistical significance. As we will see in a few examples in the next section, the observed differences are *always* highly significant (with p-values below any imaginable significance level). The point of Algorithm 3 is to predict occurrence of SH episodes on an individual basis. The results are:

(1) The minimum baseline observation period is 50 SMBG readings taken over approximately two weeks with a frequency of 3-4 readings a day. After this time each subject is classified in one of two risk groups that later use different decision-making rules;

(2) From the 6 months of data that we have we find that it is sufficient to make this group assignment once in the beginning of observation. Thus, we can assume that about every 6 months the meter would use 50 reading to reevaluate its owner's group assignment;

(3) The optimal lag of SMBG tracking is 100 to 150 readings taken with a frequency of 3-4 readings per day. In other words, the optimal decision-making criterion would be based on a computation using all 150 readings in a meter's memory. This was done to simulate the memory capacity of OneTouch Ultra. In general, good results are achieved using a lag of only 20 readings taken over a week, but a longer lag yields better prediction;

(4) The decision-making rules is based on a new computational procedure that tracks subjects Low BG Index and other related parameters using "provisional mean" computation. Special software was designed to implement this procedure and to process the data that we had available. From a programming point of view, the code needed for implementation of this procedure is only about 20 lines, which includes the computation of the LBGI;

(5) Several decision-making rules (using various parameters) were investigated. Regardless of the frequency of SMBG, these rules achieved prediction of SH within 24 hours anywhere from 43.4% with R_{ud}=1:25 to 53.4% with R_{ud}=1:7. Thus, compared to our previous result, the prediction of SH within 24 hours increased by 10%;

(6) As an optimal solution for further investigation we choose the decision-making rule that predicted 50% of SH within 24 hours and had R_{ud}=1:10. The following results refer to this optimal solution under different conditions:

(7) The optimal frequency of SMBG is 4 readings per day. If this frequency is achieved, the prediction of SH within 24 hours increases to 57.2% with the same R_{ud}=1:10. Other frequencies of SMBG are investigated and reported as well;

(8) If we extend the prediction period to 36, or 48 hours, the prediction of SH increases to 57% and 63% respectively, with the same R_{ud}=1:10;

(9) Utilizing baseline information increases substantially the prediction of SH. In fact, the 10% increase over our previous version of Algorithm 3 is entirely due to the use of baseline tracking. However, this baseline tracking is now modeled as a two-week period of self-calibration of the meter that does not use any additional input from the subject;

(10) Personal/demographic information, such as history of SH or prior HbA_{1c}, does not contribute to a better short-term prediction of SH;

(11) Raising a flag whenever there is a prolonged period of no SMBG activity is not justified. The only times when the meter would issue warning for upcoming SH would be the times of usage. This is because a major part of the prediction of SH is based on the recurrence (clustering) of very low BGs. An assessment of this recurrence is presented in an abstract (Kovatchev et al. Recurrent Hypoglycemia and Severe Hypoglycemia (SH) in T1DM Patients With History of Multiple SH) prepared for the June 2002 ADA meeting (See Appendix).

Detailed Description of the Data Processing

The meter stores SMBG readings together with the date and exact time (hour, minute, second) of each reading. Thus, in Training Data set 2 we have for each subject a certain temporal sequence of SMBG records. During the study, a total of 75,495 SMBG readings (on average 4.0±1.5 per subject per day) were downloaded from the participants' memory meters. From subjects' monthly diaries, we had the date and time of SH episodes that had occurred. Subjects reported 399 (4.7±6.0 per subject) SH episodes. Sixty-eight (80%) of the participants experienced one or more episodes of SH. These subjects did not differ from those who did not experience SH (the remaining 20% of the subjects) in terms of any of their demographic characteristics.

Pre-Processing of the Data: Special software was developed for *pre-processing of the data*. This included: (1) Assembling of the memory meter data for each subject into a continuous 6-8-month sequence of BG readings, and (2) Matching of each subject's records of SH with this sequence by date and time. The latter was performed as follows: for each SMBG reading the time (hours/minutes) until the nearest SH episode, and the time elapsed from the latest SH episode, were computed. Thus, it was possible to: (1) time 24-hour, 48-hour, etc. periods backward and forward from each SH episode, and (2) time periods between SMBG readings. Due to the nature of SH (stupor, unconsciousness), no SMBG was performed exactly at the time of SH, thus SH episodes for the purposed of Algorithm 3 do not include biochemical significant hypoglycemia that was used for Algorithm 2. The average per SH episode minimum elapsed time between SH and the nearest preceding SMBG reading was 5.2±4.1 hours; 29 SH episodes (7%) were preceded by a SMBG reading within 15 minutes. For each SH episode, we counted how many SMBG readings were performed within 24h, 36h, 48h, and 72h prior to that episode.

Computing of Baseline Risk Values and Self-Calibration: The Low BG Index for each subject is computed on his/her first SNBG readings. It was determined that the minimum number of reading required to compute a baseline LBGI is 50 taken over approximately 2 weeks. Therefore for each new meter we need to anticipate an initial two-week self-calibration period during which the meter would be scanning the overall risk for SH of its owner. After the initial period, the person is classified into one of two risk groups: Low-moderate risk (LBGI $\leq$ 3.5, LM Group) or moderate-to-high risk (LBGI > 3.5, MH Group). Our test data show that a more precise classification would not be necessary. This classification allows for different decision-making rules to be used in the LM and MH groups and raises the hit rate of the algorithm by approximately 10% as compared to its original hit rate presented in the invention disclosure.

With the test data re-calibration of the baseline risk was not necessary. Thus, we can assume that if the person does not undergo changes in treatment, re-calibration would be performed approximately every 6 months. This is consistent with the results of Algorithm 2 showing that the long-term prediction of SH is quite valid for 6 months after the initial observation period.

However, if the person experiences rapid changes in his/her glycemic control, re-calibration maybe required more frequently. The decision for re-calibration can probably be automated and based on observed increasing differences between the running risk value (see the next paragraph) and the baseline LBGI. However, the available data do not allow us to clarify this issue since the subjects that we observed did not have substantial changes in their risk for hypoglycemia.

Computing SMBG Parameters: After the pre-processing step, another piece of software was designed to compute SMBG parameters that would be used for prediction of imminent SH. This software included:

(1) Computing of a Low BG Risk value (RLO) for each BG reading that is done by the following code (here BG is measured in mg/dl, if the units are mmol/l the coefficients are different):

```
scale=(ln(bg))**1.08405 - 5.381
risk=22.765*scale*scale
if (bg_1 le 112.5) then
   RLO=risk
else
   RLO=0
endif
```

(2) For each SMBG reading with a sequential number *n*, BG(n), computing of a running value of the LBGI(n), and another statistics, SBGI(n) that is the standard deviation of the low BG risk values. These two parameters were computed with a certain lag (*k*) backwards from each SMBG reading, e.g. included that reading, BG(n), and (k-1) readings taken prior to BG(n).

(3) The computation of LBGI(n) and SBGI(n) used a *new provisional means procedure* that is based on the following recursive code:

Initial values at *n-k* (or at the max(1,n-k) to be exact in order to account for meter readings with a sequential number less than *k*):

```
LBGI(n-k)=rlo(n-k)
rlo2(n-k)=0
```

Values for any consecutive iteration *j* between *n-k* and *n*:

```
LBGI(j)=((j-1)/j)*LBGI(j-1) + (1/j)*RLO(j)
rlo2(j)=((j-1)/j)*rlo2(j-1) + (1/j)*(RLO(j)-LBGI(j))**2
```

After this cycle is completed we have the value of LBGI(n) and we compute

```
SBGI(n)=sqrt(rlo2(n))
```

Since the maximum of *n* is 150 for OneTouch Ultra meters, the search for an optimal lag *k* was performed within the range of k=10 to k=150. Although the difference in performance was not significant, the optimal lag was determined to be *k=150* (see the next section for examples).

Decision-Making Rule: At each SMBG reading the procedure decides whether to raise a flag warning for upcoming SH, or not. If the flag is raised, the procedure decides whether to bring it down. These decisions depend on three threshold parameters, α, β, γ that work as follows:

For subject at low-to-moderate risk (LM group):

```
FLAG=0
if (LBGI(n) ≥ α and SBGI(n) ≥ β) FLAG=1
if (RLO(n) ≥ (LBGI(n)+γ*SBGI(n))) FLAG=1
```

For subjects in the moderate-to-high risk group only the second if-statement is active. In other words, the flag is raised (e.g. becomes equal to 1) if both the running value of LBGI(n) and its standard deviation SBGI(n) exceed certain threshold values, and is also raised if the current value of the low BG risk RLO(n) exceeds the value of LBGI(n) plus γ standard deviations.

An heuristic explanation: The values of LBGI(n) and SBGI(n) reflect slower changes in risk for hypoglycemia – it takes a few days of SMBG to substantially change these values. Since elevated LBGI(n) means more frequent and extreme recent hypoglycemia, we can conclude that LBGI(n) and SBGI(n) reflect a persistent depletion (or lack of replenishment) of counterregulatory reserves over the course of several days. In addition, SBGI(n) is a marker of the stability of the system - a larger SBGI(n) indicates that a subjects' BG fluctuations increase and therefore the control system becomes unstable and vulnerable to extreme aberrations.

Thus, the first logical expression reflects the notion that SH occurs whenever the couterregulatory defenses are exhausted and the controls (external or internal) become unstable. The second logical expression accounts for acute changes in the low BG risk, triggering a flag whenever the current Low BG risk value suddenly becomes greater than its running average. The fact that for subjects in the moderate-to-high risk group only the second logical expression is relevant goes along with these subjects' eventual "permanent depletion" and "permanent instability" status. Since these subjects continuously run low BG values, and their BG is unstable, any acute hypoglycemic episode would be capable of triggering SH. In general, a flag for severe hypoglycemia is raised either after a period of low unstable BG, or after an acute hypoglycemic event that deviates substantially (in a risk space) from the latest running risk average (that maybe already high). It follows that SH episodes that are not preceded by any of these warning signs will remain unaccounted for by this algorithm. Below we present a sample output that illustrates the action of Algorithm 3 for several subjects:

ID	BG	SH	FLAG	TIME
135	70	.00	.00	53.75
135	77	.00	.00	41.09
135	124	.00	.00	35.02
135	51	.00	1.00	30.44
135	50	.00	1.00	14.72
135	66	.00	< 24h	10.60
135	49	.00	1.00	8.30
135		1.00		
...				
135	97	.00	.00	140.05
135	130	.00	.00	25.17
135	59	.00	.00	20.20
135	76	.00	.00	5.23
135	41	.00	1.00	.62
135		1.00		
219	200	.00	.00	40.72
219	64	.00	.00	37.88
219	43	.00	1.00	28.73
219	225	.00	.00	16.22
219	38	.00	1.00	11.18
219	43	.00	< 24h	10.87
219	75	.00	< 24h	4.52
219		1.00		
222	156	.00	.00	19.08
222	176	.00	.00	13.23
222	83	.00	.00	9.72
222	66	.00	.00	7.83
222	42	.00	1.00	4.75
222		1.00		
223	228	.00	.00	18.80
223	149	.00	.00	14.15
223	41	.00	1.00	5.85
223		1.00		
223	110	.00	< 24H	6.00
223		1.00		

Here $\alpha=5$, $\beta=7.5$, $\gamma=1.5$

For subject # 135 the first flag is raised about 30 hours prior to SH and stays up for the next reading whish is taken 16 hours later and 14 hour prior to SH. This latter reading and the two readings that follow are within 24 hours prior to SH, so we consider this episode to be predicted. Nevertheless, the flag is brought up again about 8 hours and 20 minutes prior to SH.

A second SH episode for this subject is flagged 36 minutes in advance.

This subject gets two warnings – approximately 28.7 and 11.2 hours prior to this SH episode.

This subject gets a warnings approximately 4 hours and 45 minutes prior to this SH episode.

This subject experienced two recurrent SH episodes within 12 hours. The flag is raised approximately 6 hours before the first episode and therefore we consider both episodes to be in the predicted risk high-risk 24-hour time period.

Each line of this output presents an SMBG reading, or an SH episode (without a reading). ID is subject's ID number, BG is BG level in mg/dl, SH=1 whenever SH episode occurs. FLAG=1 if Algorithm 3 decides to raise the flag; TIME is the time to the nearest SH episode in hours.

Optimizing the of Lag of the Provisional Means Procedure: In a prior publication we have reported that in the period 48 to 24 hours before SH the average BG level decreased and the BG variance increased. In the 24-hour period immediately preceding SH average BG level dropped further, the variance of BG continued to increase, and there was a sharp increase in the LBGI. In the 24-hour period following SH, average BG level normalized, however the BG variance remained greatly increased. Both the average BG and its variance returned to baseline levels within 48 hours after SH (see Kovatchev et al. Episodes of Severe Hypoglycemia in Type 1 Diabetes are Preceded, and Followed, within 48 Hours by Measurable Disturbances in Blood Glucose. J of Clinical Endocrinology and Metabolism, 85: 4287-4292, 2000). We now use these observations to optimize the lag of the provisional means procedure, *k*, employed by Algorithm 3 on the basis of the deviations in the average values of LBGI(n) and SBGI(n) observed within 24 hour prior to SH. In short, the lag for computing LBGI(n) and SBGI(n) was chosen to maximize the difference that these measures display within 24 hours prior to SH compared to the rest of the study, excluding periods immediately after SH when the system is out of balance. The optimal lag was found to be *k*=150. Table 6 presents the means of LBGI(n) and SBGI(n) for several values of the parameter k. and for both subject groups, low-moderate risk and moderate-high risk. It is evident that the difference between various values of *k* is not great, thus in a practical application any value of $k \geq 10$ would be appropriate. However, based on the current data we would recommend k=150, and all further computations use this lag. This recommendation is also based on the reduced variance in LBGI(n) and SBGI(n) at larger lag values, that is reflected by larger t-values below:

Table 6A: LBGI(n) within 24 hour prior to SH vs. the rest of the time for different lags:

LBGI	Low-moderate risk (LM Group)				Moderate-high risk (MH Group)			
	24h prior SH	Rest of the time	t	p	24h prior SH	Rest of the time	t	p
k=10	4.99	3.32	9.2	<.000...	6.94	5.28	11.0	<.000...
k=20	4.73	3.34	9.4	<.000...	6.56	5.32	11.0	<.000...
k=30	4.54	3.35	9.1	<.000...	6.50	5.34	11.3	<.000...
k=50	4.53	3.34	9.8	<.000...	6.54	5.36	11.7	<.000...
k=100	4.45	3.29	11.2	<.000...	6.53	5.40	12.3	<.000...
k=150*	4.46	3.26	12.1	<.000...	6.56	5.41	12.9	<.000...

Table 6B: SBGI(n) within 24 hour prior to SH vs. the rest of the time for different lags:

SBGI	Low-moderate risk (LM Group)				Moderate-high risk (MH Group)			
	24h prior SH	Rest of the time	t	p	24h prior SH	Rest of the time	t	p
k=10	7.50	5.35	9.0	<.000...	9.64	7.78	11.4	<.000...
k=20	8.31	6.28	10.1	<.000...	10.36	8.91	10.7	<.000...
k=30	8.50	6.68	9.5	<.000...	10.82	9.38	11.5	<.000...
k=50	8.78	7.04	10.0	<.000...	11.30	9.80	12.3	<.000...
k=100	9.19	7.33	11.7	<.000...	11.69	10.18	13.5	<.000...
k=150*	9.42	7.42	13.7	<.000...	11.88	10.33	15.1	<.000...

* Optimal solution

As seen in Tables 6A and 6B both LBGI and SBGI become highly significantly elevated in the 24-hour periods preceding SH. Thus, one is tempted to run a direct discriminant or logistic model to predict upcoming SH. Unfortunately such standard statistics don't work very well, although both models are highly statistically significant. The discriminant model (that worked better than logistic regression) predicted correctly 52.6% of upcoming SH episodes. However, its flag-up to flag-down ratio was quite poor - R_{ud}=1:4. Therefore this model was biased towards the larger amount of data points, a bias that is to be expected in any statistical procedure. Consequently, we had to employ the decision-making rule presented above.

Accuracy of Prediction of Severe Hypoglycemia

Optimization of the Threshold Parameters α, β, and γ: Below we present a detailed account of the predictive power of Algorithm 3 using various combinations of its threshold parameters α, β, and γ. Since the relationship between these parameters and the desired outcome (high prediction of SH and minimal ratio R_{ud}) is quite complex, the optimization procedure that we used did not reach a single solution. Also, it seems that there is no need for a single solution either. It is probably a business, rather than mathematical decision, what would be an acceptable % of prediction of SH, given a "flag up" to "flag down" ratio. Thus, we do not claim that any of the presented below solution is optimal. However, in order to explore this subject further, we accept that a 50% prediction of future SH with R_{ud} = 1:10 is a base for investigating other than 24 hours prediction periods as well as various requirements for the number of SMBG readings per day required for a better risk profile.

Table 7 presents the performance of Algorithm 3 at several combinations of the values of α, β, and γ that are representative for the relationship between the percentage of predicted SH (hit rate) and the ratio R_{ud} which we could call "annoyance index." Table 7 also includes the average total time (in days) per subject spent in alert vs. no-alert status during the study, i.e. the summary result from the alternating process of warning – no-warning periods that a subject would experience using this algorithm which illustrates the meaning of the ratio R_{ud}.

Table 7: Prediction of SH: Hits, Annoyance Index, and Average Times:

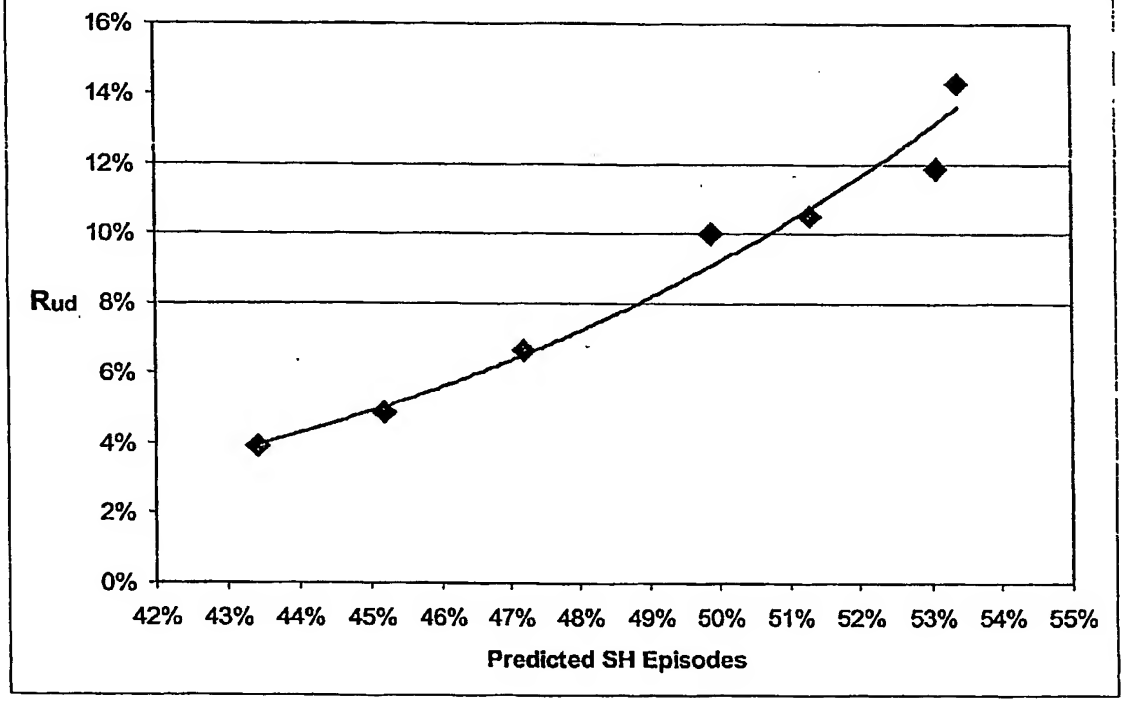
					Total for study (days)	
α	β	γ	% Hit	R_{ud}	Flag up	Flag Down
6.4	8.2	1.5	43.4	1 : 25	7.8	198.9
6.0	7.5	1.5	45.2	1 : 20	9.6	197.3
5.5	7.5	1.5	47.2	1 : 15	12.9	194.1
5.0	7.5	1.5	49.9	1 : 10	19.0	190.1
5.0	7.5	1.3	51.3	1 : 9.5	19.5	185.7
4.9	7.0	1.2	53.1	1 : 8.4	21.6	182.0
4.8	7.0	1.2	53.4	1 : 7	25.5	178.2

The highlighted solution is used for all further analyses. Given that the participants in this study experienced 4.7 SH episodes on average, 19 days of high-alert periods seem to be acceptable,

if these alerts would prevent 50% of SH. In addition, high-alert periods tend to come in clusters. Therefore we can assume that in practice, long and relatively calm periods will alternate with a few days of high-risk warnings. The last line in Table 7 presents a solution with a R_{ud} = 1:7, which is equivalent to the solution presented in the invention disclosure. However, the current solution has almost 10% higher hit rate, 53.4% compared to 44% in our previous algorithm. When the hit rate is comparable to our previous algorithm, the annoyance ratio is below 1:20, i.e. three times better.

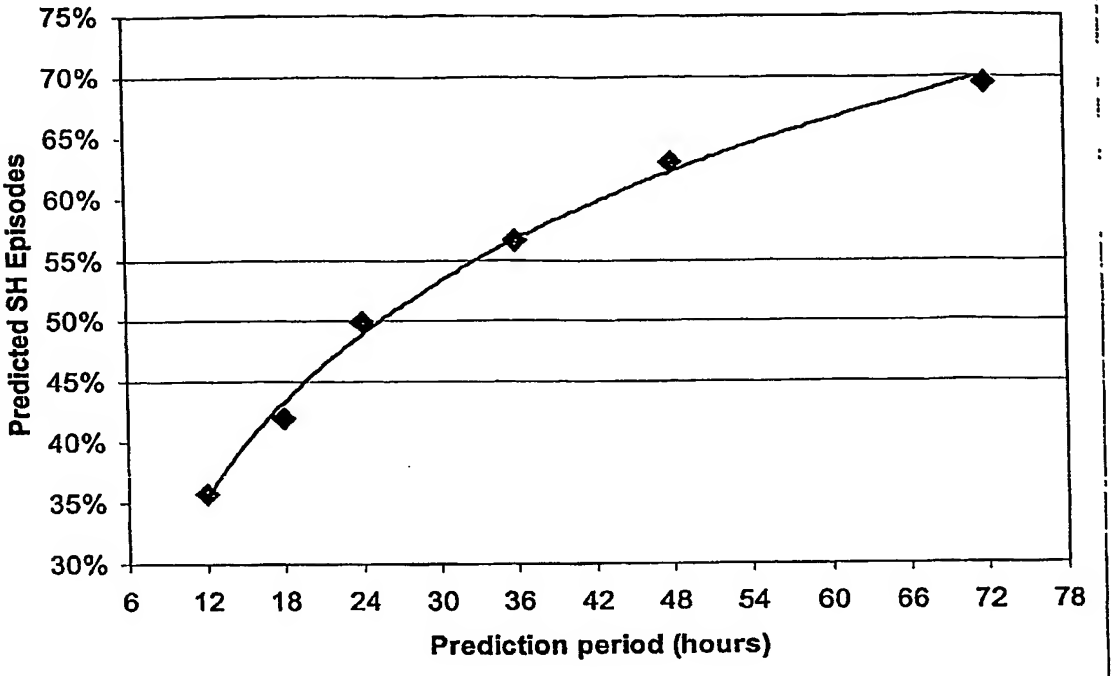
Figure 7 presents the smoothed dependence between the hit rate and the ratio R_{ud} expressed in percentage. It is evident that the ratio between "flag up" and "flag down" increases rapidly when the hit rate of Algorithm 3 increases. Thus, given these data it maybe unjustified to pursue parameter combinations resulting in a higher than 50% hit rate:

Figure 7: Hit Rate vs. Annoyance Index



Alternative Prediction Periods: In the beginning of the description of Algorithm 3 we made the basic assumption that an SH episode would be considered predicted if the flag is raised within the 24-hour period of time preceding this episode. This assumption resulted in the hit rates reported in the previous section. We will now present computations of the hit rate based on other prediction periods ranging from 12 to 72 hours. Throughout this experiment the parameters α, β, and γ remain fixed at 5.0, 7.5, and 1.5 respectively, i.e. at their values in the solution highlighted of Table 7. Therefore the flag-up rate remains the same as in this solution with R_{ud}=1:10, and only the hit rate changes since we change the definition of a hit. Figure 8 presents the dependence between the prediction period and the corresponding hit rate.

Figure 8: Prediction Period vs. Hit Rate (Rud=10)



It is evident that the hit rate increases rapidly with the increase of the prediction period to about 24 hours and then the increase of the hit rate gradually slows down. Therefore, we can conclude that 24 hours ahead is an optimal and a reasonable forecast period.

Optimal Number of SMBG Readings Per Day. Finally we experiment with the requirement of how many readings per day are needed in order to produce an optimal forecast of SH.

As we said in the beginning, all reported SH episodes were 399. Of these episodes 343 had any SMBG reading available in the preceding 24 hours (additional 3 episodes had any reading within the preceding 48 hours and additional 4 episodes had any reading within the preceding 72 hours). It follows that more than 50 SH episodes (14%) did not have any reasonable preceding SMBG reading that would help with their prediction. The 343 episodes that had at least one prior SMBG reading within 24 hours were used for the computation of the hit rates in the previous section. The other episodes were naturally excluded from the computation.

Further analysis shows that the hit rate increases rapidly with the number of readings taken before an SH episode. However, if we impose a strict requirement for a certain number of readings to be available in order to consider an SH episode, we see that the number of SH episodes that meet this requirement rapidly decreases (Table 8). This is due to subjects' non-compliance with the study requirements and is maybe a good reason to incorporate in future meters some sort of a warning message that Algorithm 3 will not be useful and would be switched off if there are no SMBG readings taken at an appropriate rate.

Table 8 presents the number of SH episodes that had available certain number of preceding SMBG readings and the hit rate of Algorithm 3 for these episodes. The highlighted row of the table contains the optimal solution from Table 7 that was used as a base for all subsequent computations. All hit rates are given in terms of a 24-hour prediction period, i.e. flag within the 24 hours preceding SH. We can conclude that with an increased subjects' compliance the accuracy of Algorithm 3 in prediction SH would increase substantially. With 5 SMBG reading per day the accuracy is up 10% from its base of 50% hit:

Table 8: Performance of Algorithm 3, Given a Certain Number of Prior SMBG Readings

Number of Preceding SMBG Readings	SH episodes that satisfy the requirement in column 1 (% of total number of SH)	Hit Rate
At least 1 within 24 hours	343 (86%)	49.9%
At least 3 within 24 hours	260 (65%)	54.2%
At least 4 within 24 hours	180 (45%)	57.2%
At least 5 within 24 hours	103 (26%)	64.1%
At least 4 within 36 hours	268 (67%)	52.6%
At least 5 within 36 hours	205 (51%)	54.6%
At least 6 within 36 hours	146 (37%)	60.3%
At least 7 within 36 hours	107 (27%)	60.7%
At least 6 within 48 hours	227 (57%)	53.3%
At least 7 within 48 hours	187 (47%)	54.0%
At least 8 within 48 hours	143 (36%)	55.9%
At least 9 within 48 hours	107 (27%)	59.8%

Other Potential Enhancements That Were Tested

The attempts to increase the predictive power of Algorithm 3 by inclusion of external parameters, such as number of SH episodes in the previous year, or baseline HbA1c were unsuccessful. Evidently, the short-term prediction of SH is mainly dependent on current or recent events. However, a limitation of this study is that all participating subjects had a history of ≥ 2 SH episodes in the previous year.

Finally, we tested whether an alert for SH should be issued if an increased risk for hypoglycemia is detected and there is no SMBG for certain period of time. This was done in an attempt to predict at least some of the SH episodes that were not preceded by any SMBG readings. This was not successful, generating predominantly false alarms. This result comes as an additional confirmation of the importance of compliance with an SMBG protocol comprised of sufficiently frequent SMBG readings.

Appendix: Abstract Accepted for the 2002 Meeting of the American Diabetes Association

Recurrent Hypoglycemia and Severe Hypoglycemia (SH) in T1DM Patients With History of Multiple SH

Boris Kovatchev, Daniel Cox, Linda Gonder-Frederick, William Clarke, University of Virginia, Charlottesville, VA.

This study evaluates the frequency of recurrent hypoglycemia and SH (defined as stupor or unconsciousness that preclude self-treatment) following a low blood glucose (BG<3.9mmol/l) episode.

Eighty-five patients (41 female) with T1DM and history of >2 episodes of SH in the last year performed SMBG 3-5 times per day for 6 to 8 months and recorded in diaries any SH episodes by date and time. Subjects' average age was 44±10 years, duration of diabetes 26±11 years, HbA_{1c} 7.7±1.1%.

All SMBG readings (n=75,495) were merged by date and time with subjects' SH episodes (n=399; SH events generally do not have a corresponding SMBG reading). For each SMBG reading, or SH episode, the elapsed time since the nearest previous low BG (<3.9mmol/l) was computed. The table below presents the percentage of readings in 3 hypoglycemic ranges: BG<1.9mmol/l, 1.9-2.8mmol/l, and 2.8-3.9mmol/l, as well as the percentage of SH episodes, that were preceded by a Low BG reading (BG<3.9mmol/l) within 24 hours, 24-48 hours, 48-72 hours, and more than 72 hours. The last column presents Runs tests rejecting the hypotheses that the days containing low BG readings (or SH episodes) are randomly distributed across time. The negative Z-values of the tests show "clustering" of days with and without hypoglycemic readings or SH episodes.

Percentage of hypoglycemia/SH preceded by a low BG:

BG	< 24h.	24-48h.	48-72h.	> 72h.	Runs Test	
					Z	p-level
<3.9 mmol/l	50%	21%	10%	19%	-13.6	<.0001
2.8-3.9 mmol/l	52%	20%	10%	18%	-18.3	<.0001
1.9-2.8 mmol/l	55%	20%	7%	18%	-14.7	<.0001
SH	64%	11%	6%	19%	-11.1	<.0001

We conclude that more than half of all hypoglycemic SMBG readings and approximately 2/3rds of all SH episodes, are preceded by at least one hypoglycemic reading within the previous 24 hours. In addition, hypoglycemic events tend to appear in clusters. Thus, an initial hypoglycemic episode may be a warning sign for upcoming recurrent hypoglycemia.